Differential Geometry I, Winter Semester 2024–25, HU Berlin

Chris Wendl

Contents

These notes will be expanded gradually over the course of the semester. If you notice any typos or mathematical errors, please send e-mail about them to `wendl@math.hu-berlin.de` and they will be corrected.

While the notes are written in English, I make an effort to include the German translations (*geschrieben in dieser Schriftart*) of important terms wherever they are introduced. I will occasionally omit these translations in cases where the English and German words are identical, or if the word has already appeared before with its translation in a different context (e.g. the word "smooth" needs to be defined many times in different contexts, and its German translation is always the same), and also in cases where I can't reliably figure out what the German word is. The latter will happen more often as the course goes on, because the deeper one gets into advanced mathematics, the harder it becomes to find authoritative German sources for clarifying the terminology (and I am not linguistically qualified to invent terms in German myself).

Most recent update: **December 11, 2024**

1. Introduction

Before diving in with definitions, theorems and proofs, I want to give an informal taste of what differential geometry is all about. The word "informal" means, in this case, that you should try not to worry too much about the precise definitions or rigorous arguments behind what we are discussing, but focus instead on the big picture. Before the first lecture is finished, I will revert to being a proper mathematician and give some actual definitions.

1.1. A foretaste of Riemannian geometry. Let's assume for the moment that we all understand what a "smooth surface" is, e.g. you can picture it as a subset[1] of $\mathbb{R}^3$ such that every point has a neighborhood parametrized by some injective[2] C^∞-map

$$\mathbb{R}^2 \overset{\text{open}}{\supset} \mathcal{U} \hookrightarrow \mathbb{R}^3.$$

With this understood, assume

$$\Sigma \subset \mathbb{R}^3$$

is a smooth surface.

1.1.1. *Distances and geodesics.* We could view Σ as a metric space by defining the distance between two points $x, y \in \Sigma$ via the Euclidean metric, but this is not necessarily the most natural thing to do. A more natural notion of distance in the surface Σ would be one that tells you something about the actual distance that an ant has to travel if it walks a path *along the surface* between x and y. If that path is parametrized by a smooth map $\gamma : [a, b] \to \mathbb{R}^3$ satisfying $\gamma([a, b]) \subset \Sigma$, $\gamma(a) = x$ and $\gamma(b) = y$, then the distance travelled is

$$\ell(\gamma) := \int_a^b |\dot{\gamma}(t)|\, dt = \int_a^b \sqrt{\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle}\, dt, \tag{1.1}$$

where $\dot{\gamma}(t)$ denotes the time derivative of $\gamma(t)$, $\langle v, w \rangle$ denotes the Euclidean inner product of two vectors $v, w \in \mathbb{R}^3$, and $|v| := \sqrt{\langle v, v \rangle}$ denotes the Euclidean norm. If we denote by $\mathcal{P}(x, y)$ the set of all smooth paths in Σ connecting x to y, then a natural notion of distance on Σ can now be defined by

$$d(x, y) := \inf_{\gamma \in \mathcal{P}(x,y)} \ell(\gamma). \tag{1.2}$$

The infimum needs to be taken since, in general, there are many distinct paths from x to y that will have different lengths. In principle we are interested in the *shortest* such path, though it is not obvious in general whether such a shortest path must exist:

QUESTION 1.1. *Given a smooth surface Σ and two distinct points $x, y \in \Sigma$, does there exist a smooth path on Σ from x to y that has the* shortest *possible length? Is it unique?*

We will see later in this semester that the answer to both questions is always yes if x and y are close enough to each other, and the shortest path can then be characterized by a second-order ordinary differential equation. Such a path is called a **geodesic** (*Geodäte* or *geodätische Linie*), and it serves as the best possible substitute for a "straight line" on Σ, even in cases where no actual *straight* paths on Σ exist. The canonical example you should picture is the unit sphere $\Sigma := S^2 \subset \mathbb{R}^3$, whose geodesics are the so-called *great circles*, namely the subsets $S^2 \cap P$ defined via 2-dimensional linear subspaces $P \subset \mathbb{R}^3$. These are the paths that all airplanes would traverse

[1]We will soon improve this definition so that surfaces do not need to be regarded as subsets of $\mathbb{R}^3$. In fact, there are some important examples of surfaces that *cannot* be embedded in $\mathbb{R}^3$; a famous example is the Klein bottle, see https://en.wikipedia.org/wiki/Klein_bottle.

[2]We will need to add a condition concerning the derivative of the map $\mathcal{U} \hookrightarrow \mathbb{R}^3$ before this becomes an adequate definition, but let's worry about that later.

along the Earth if there were no additional factors such as weather conditions or no-fly zones to consider.

1.1.2. *Angles, isometries, and curvature.* The fundamental piece of data that makes the above definition of distance on Σ possible is the Euclidean inner product $\langle\ ,\ \rangle$. In fact, $\langle\ ,\ \rangle$ contains strictly more information than is actually needed for defining distances on Σ; if you look again at the formula (1.1), you'll notice that it doesn't really require knowing what $\langle v, w\rangle$ is for every $v, w \in \mathbb{R}^3$, but is already well-defined if we know how to define this for every pair of vectors v, w that are *tangent* to Σ at any given point. (Indeed, $\dot{\gamma}(t) \in \mathbb{R}^3$ is always tangent to Σ at $\gamma(t)$.) In fact, it would suffice to know what $\langle v, v\rangle$ is for every individual tangent vector v, but knowing $\langle v, w\rangle$ for two distinct vectors provides some additional information that is of geometric interest: it allows us to compute the *angle* between any two tangent vectors. Indeed, the angle θ between two vectors $v, w \in \mathbb{R}^3$ can always be deduced from the formula

$$\langle v, w\rangle = |v| \cdot |w| \cdot \cos\theta.$$

We can therefore define not only the length of any smooth path along Σ, but also the angle between two smooth paths wherever they intersect. This information makes Σ into what we will later call a (2-dimensional) **Riemannian manifold** (*Riemannsche Mannigfaltigkeit*), and the restriction of the inner product to the tangent spaces on Σ, which determines all lengths and angles, is called a **Riemannian metric** (*Riemannsche Metrik*).[3]

Here is a natural question one can ask about Riemannian manifolds. Suppose $\Sigma_1, \Sigma_2 \subset \mathbb{R}^3$ are two smooth surfaces, and $\varphi : \Sigma_1 \to \Sigma_2$ is a smooth bijective map between them whose inverse is also smooth.[4] We call φ in this case a **diffeomorphism** (*Diffeomorphismus*), and say that Σ_1 and Σ_2 are **diffeomorphic** (*diffeomorph*). We say that φ is additionally an **isometry** (*Isometrie*) if it preserves all distances and angles, and in this case, Σ_1 and Σ_2 are said to be **isometric** (*isometrisch*).

QUESTION 1.2. *Given two diffeomorphic surfaces, how can we measure whether they are isometric?*

In simple examples, it is often easy to recognize when two surfaces are diffeomorphic: an example is shown in Figures 1 and 2, where we can compare the standard unit sphere $S^2 \subset \mathbb{R}^3$ with a "nonstandard" embedding of S^2 into $\mathbb{R}^3$ that elongates a portion of the sphere into something more closely resembling a cylinder. It is surely not hard to imagine that these two surfaces in $\mathbb{R}^3$ are diffeomorphic; writing down an explicit example of a diffeomorphism would be a pain in the neck, but we will content ourselves with the intuitive understanding that in the process of "stretching" the standard sphere into its nonstandard counterpart, one could if necessary come up with a smooth bijection between the two. The much deeper observation is that they are *not* isometric, and we will need to develop some technology before we can prove this rigorously. One of the key ideas behind the proof is shown in Figures 1 and 2: on any surface Σ, one can draw a closed piecewise-smooth path along Σ, choose a starting point p_0 on the path and a tangent vector v_0 at p_0, then translate the vector v_0 along the path via a process known as **parallel transport**. We will have to give a careful definition later of what is meant by parallel transport, but Figures 1 and 2 will hopefully give you some intuition about this. The interesting question is now: if we parallel transport the

[3] Caution: there is a potential for confusion in this terminology, because a Riemannian metric is not a particular kind of metric in the sense of metric spaces, though it does determine one via formulas such as (1.2). A Riemannian metric carries strictly more information, since it determines angles in addition to distances.

[4] For the purposes of this discussion, you may assume that a function on a smooth surface $\Sigma \subset \mathbb{R}^3$ is smooth if it can be extended to a smooth function on a neighborhood of Σ; the latter notion is familiar from your first-year Analysis class since the neighborhood is an open subset of $\mathbb{R}^3$. We will later give an equivalent but more elegant definition of smoothness for functions on manifolds.

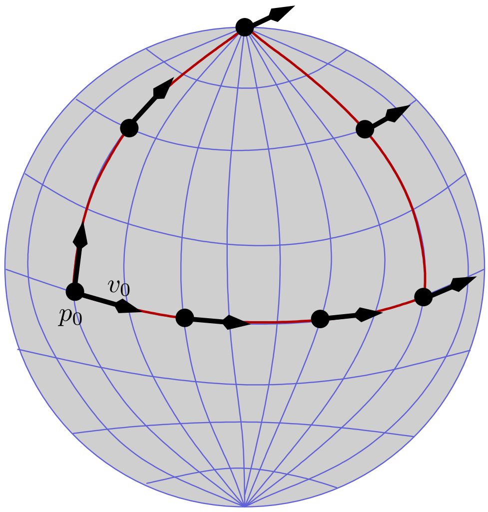


FIGURE 1. The "round" sphere $S^2 \subset \mathbb{R}^3$. Parallel transport of a vector along a closed path leads to a different vector upon return.

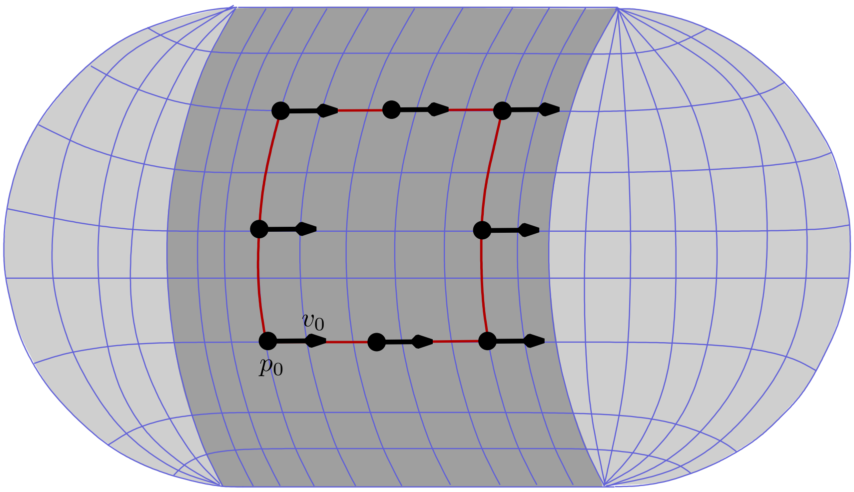


FIGURE 2. A different embedding of S^2 in $\mathbb{R}^3$, so that the darkly shaded region is locally flat. Parallel transport of a vector around a closed path in this region always leads back to the same initial vector.

vector v_0 once around our chosen closed path, does it return to the same starting vector? As you can see in the pictures, the answer is no for the triangular path in Figure 1, but yes for the rectangular path in Figure 2. It will turn out that this observation encodes a fundamental difference between these two Riemannian manifolds: the standard sphere has positive **curvature** (*Krümmung*) at every point, but the elongated sphere does not—if fact, the surface in Figure 2 has zero curvature everywhere on the elongated region where our rectangle is drawn.

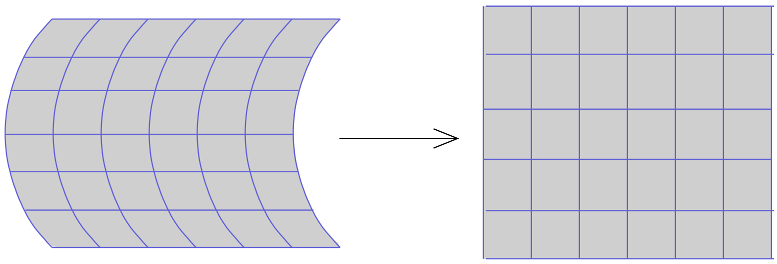


FIGURE 3. A piece of a cylinder can be flattened to a plane without changing any lengths or angles on the surface.

A major portion of the second half of this semester will be devoted to the precise definition of curvature and its important properties. One of these is that it completely characterizes the notion of *local flatness*:

QUESTION 1.3. *Given a smooth surface $\Sigma \subset \mathbb{R}^3$ and a point $p \in \Sigma$, does p have any neighborhood that is isometric to an open subset of the "flat" surface $\mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$?*

A surface $\Sigma \subset \mathbb{R}^3$ is called **locally flat** (*lokal flach*) if the answer to Question 1.3 is yes for every point $p \in \Sigma$. Figure 3 shows an example of a surface that is locally flat, even though it does not look flat in the picture: you know it is locally flat because you know that an ordinary piece of paper can be bent into this cylindrical shape without breaking or stretching it. This is *not* true of the standard unit sphere in $\mathbb{R}^3$. Perhaps you've never held in your hand a piece of paper that's shaped like part of a globe[5], but you can surely imagine that if you did, you could never make it *flat* without breaking or stretching it. This is another symptom of the positive curvature of the round sphere.[6] By contrast, the cylindrical surface in Figure 3 has zero curvature everywhere. The statement that a cylinder is in some sense "not curved" may seem jarring at first, but you'll get used to it: the point is that the quantity we're calling curvature should depend only on the Riemannian metric, and not on the specific way we've chosen to embed our Riemannian manifold in $\mathbb{R}^3$. If two surfaces are isometric, then their curvatures at corresponding points will always be the same.

The positive curvature of the round sphere is not unrelated to the fact that the angles of the "triangle" in Figure 1 add up to considerably more than 180 degrees. We will later also see examples of surfaces with *negative* curvature: the basic picture to have in mind is the shape of a *saddle*. In these surfaces, the angles in a triangle will add up to *less* than 180 degrees. The elongated sphere in Figure 2 has zero curvature in the shaded region, but not everywhere; since it is diffeomorphic to S^2, one could reinterpret this as the statement that S^2 admits a Riemannian metric that is locally flat in some region. That is not a deep or surprising statement, as *every* Riemannian metric on an arbitrary manifold can in fact be modified to make it flat in some small region. A more interesting question is whether it can be modified to make it locally flat *everywhere*, like the cylindrical surface in Figure 3. Let us take this opportunity to state a standard corollary of a rather deep theorem:

THEOREM. *There is no Riemannian metric on the sphere S^2 that is everywhere locally flat.*

[5] If you know where to buy one, please let me know!

[6] This is also the mathematical reason why it is impossible to create a flat map of the Earth without distorting distances and angles in some regions.

This will follow from the beautiful *Gauss-Bonnet theorem* for surfaces, to be proved near the end of this semester. It relates the integral of the curvature over a compact surface to a topological quantity, its *Euler characteristic*, which in the case of S^2 is positive. This is the reason why Figure 2 could not have been drawn so that *every* part of the sphere had zero curvature. We will also use a variant of this theorem to understand what the various observations above about sums of angles of triangles have to do with curvature.

1.1.3. *Spacetime as a pseudo-Riemannian 4-manifold.* Differential geometry is not only about surfaces, and it also plays an important role in subjects that cannot accurately be called "pure" mathematics. This is true especially in several areas of theoretical physics, the most famous of which is Einstein's theory of gravitation, known as the general theory of relativity *(allgemeine Relativitätstheorie)*. We will not directly discuss gravitation in this course, but several of the mathematical concepts we will cover are essential for understanding Einstein's picture of the universe.

The paradigm introduced by Einstein for an understanding of space and time can be summarized as follows:

(1) There are three spatial dimensions, but time adds a fourth. Locally, an "event" occurring in a particular place at a particular time thus requires four coordinates for its description, defining a point in $\mathbb{R}^4$.
(2) The picture in item (1) is only local, i.e. it is sufficient for describing interactions between events on a small or medium scale, but one should not assume that the set of all events in the universe (known as **spacetime** or *Raumzeit*) is in bijective correspondence with $\mathbb{R}^4$. In general, spacetime could be any smooth 4-dimensional manifold.
(3) Spacetime is endowed with a (pseudo-)Riemannian metric, which determines a notion of geodesics. In the absence of forces other than gravity, all objects move along geodesics in spacetime.
(4) The presence of mass affects the curvature of spacetime and thus changes the geodesics. A precise relationship between mass and curvature is given by the Einstein equation, the fundamental field equation of general relativity.

In this paradigm, gravity is not a force: it is just a geometric effect produced by the interaction between mass and curvature. In other words, the reason a brick falls toward the Earth if you drop it is that as soon as you let go, it starts following a geodesic in spacetime, and the Earth's mass causes curvature that determines the shape of that geodesic: moving forward in time while moving closer to the Earth in space.

I should say a word about the appearance of the prefix "pseudo-" in the above paradigm, which places Einstein's theory slightly outside the realm of standard Riemannian geometry. As sketched above, a Riemannian metric on a manifold M is a choice for each point $p \in M$ of an inner product on the space of tangent vectors to M at p. As you know, an inner product $\langle\ ,\ \rangle$ on a real vector space V is a positive-definite bilinear form, implying in particular that it is

- **symmetric**: $\langle v, w\rangle = \langle w, v\rangle$ for all $v, w \in V$;
- **nondegenerate**: For every $v \in V\backslash\{0\}$, there exists $w \in V$ such that $\langle v, w\rangle \neq 0$.

To define a pseudo-Riemannian metric on M, one adopts these two assumptions for the inner product $\langle\ ,\ \rangle$ on the space of tangent vectors at every point $p \in M$, but without assuming any positivity, i.e. we do not require $\langle v, v\rangle$ to be positive whenever $v \neq 0$. The classification of quadratic forms (or equivalently the spectral theorem for symmetric linear maps) implies that any n-dimensional vector space V with a symmetric nondegenerate bilinear form $\langle\ ,\ \rangle$ can be split into two orthogonal (with respect to $\langle\ ,\ \rangle$) subspaces

$$V = V_+ \oplus V_-$$

such that $\langle\ ,\ \rangle$ is positive-definite on V_+ and negative-definite on V_-. (Note that if both subspaces are nontrivial, then there always also exist nonzero vectors $v \in V$ such that $\langle v, v\rangle$ is zero—this does not contradict nondegeneracy!) The pseudo-Riemannian metrics used in general relativity have the property that on every tangent space, $\dim V_+ = 3$ and $\dim V_- = 1$.[7] Pseudo-Riemannian metrics with this property are also sometimes called **Lorentzian metrics**, and said to have **Lorentz signature**.

The canonical example of a Lorentzian inner product is what is called the **Minkowski metric** on $\mathbb{R}^4$: we define it by

$$\langle x, y\rangle = -x^0 y^0 + \sum_{j=1}^{3} x^j y^j, \tag{1.3}$$

where we are following the physicists' convention of labeling vectors $v \in \mathbb{R}^4$ by their coordinates v^μ with $\mu = 0, 1, 2, 3$. It is actually crucial for Einstein's theory that the metric on spacetime is *not* positive-definite, because the Lorentzian signature is precisely what produces qualitative physical distinctions between the three spatial dimensions and the fourth one, time. In the convention used above to write down the Minkowski metric, time is labelled as the zeroth coordinate, and is thus distinguished by the minus sign appearing in (1.3). More generally, a vector v in a vector space V with a Lorentzian inner product $\langle\ ,\ \rangle$ is called **time-like** if $\langle v, v\rangle < 0$, **space-like** if $\langle v, v\rangle > 0$, and **light-like** if $\langle v, v\rangle = 0$. With a bit of linear algebra, one can see that the set of all space-like vectors is connected, but the set of vectors that are time-like or light-like splits into two connected components, which we think of as representing motion *forward* or *backward in time*. Similarly, on a Lorentzian manifold, a geodesic can be either time-like, light-like or space-like, and in the first two categories one can distinguish between parametrizations of the geodesic that are oriented forward or backward in time, while for space-like geodesics there is no such distinction. The physical significance of these observations is the following: in general relativity, all particles with mass travel through spacetime along time-like geodesics, while particles with no mass travel along light-like geodesics—the latter are the particles that observers perceive as travelling at the speed of light. As far as we know, *nothing* travels along space-like geodesics, which is equivalent to saying that nothing travels faster than light. According to the geometry of spacetime, anything that *could* do this would also sometimes be observed to travel backward in time. Naturally, the non-existence of such particles according to the known laws of physics has not stopped physicists from giving them a name—tachyons—and they are mentioned frequently in science fiction, as a clearly necessary ingredient in time travel.

While we will probably not say anything further about general relativity in this course, we will prove some results about pseudo-Riemannian manifolds, and will try to avoid assuming that inner products are positive-definite unless that assumption is absolutely necessary.

1.1.4. *Gauge theory.* To round out this motivational introduction, I want to mention briefly another area of physics beyond general relativity where differential geometry plays a key role. The last half-century has witnessed intense and fruitful interactions between geometry and quantum field theory (on which the theory of elementary particles is based), along with its more exotic and controversial cousin, string theory. Each of the classical fields underlying the various types of elementary particles can be described mathematically as a geometric object, namely a *section* of a *smooth fiber bundle*. The particles that mediate the electromagnetic, strong and weak nuclear forces, in particular, are described via so-called *gauge fields*, which are known to mathematicians as *connections*: these are a fundamental piece of geometric data on a fiber bundle, analogous to the Lorentzian metrics on the spacetime manifold of general relativity. This subject as a whole is known as *gauge theory*, a term which means slightly different things in the two fields: physicists

[7] Or possibly the other way around—the literature is not unanimous on this convention.

understand it as the basis of their understanding of the forces of nature, while for mathematicians, it is a powerful framework for developing geometric and topological invariants based on spaces of solutions to nonlinear PDEs. In the big picture, gauge theory is both, and it has served as one of the most exciting sources of interactions between theoretical physics and pure mathematics during the past few decades. We will lay a few of the basic foundations for this subject via the study of vector bundles in the second half of this semester.

1.2. Charts and transition maps. We now begin the study of differential geometry in earnest.

The fundamental objects of study in this subject are called *smooth, finite-dimensional manifolds*. We will spend most of the first two lectures explaining the definition of this term and giving some basic examples.

We start with the intuition that a 1-dimensional manifold is what you have previously called a "curve" (*Kurve*), and a 2-dimensional manifold is a "surface" (*Fläche*). For arbitrary $n \in \mathbb{N}$, an elementary example of an n-dimensional manifold will be the so-called **n-sphere**

$$S^n := \left\{ x \in \mathbb{R}^{n+1} \;\middle|\; |x| = 1 \right\},$$

where $|\cdot|$ again denotes the Euclidean norm. The word "sphere" (*Sphäre*) on its own normally refers to the familiar case $n = 2$, though it can also refer to the general case if the value of n is clear from context. The 1-sphere has been known to you since Kindergarten under a different name: the **circle** (*Kreis*). Let us examine this example a bit more closely, and clarify in particular the following point: S^1 is defined as a subset of $\mathbb{R}^2$, so why do we consider it a "one-dimensional" object?

The answer can be explained via an intelligent choice of coordinates. Consider the standard polar coordinates (r, θ) on $\mathbb{R}^2$, which are related to the Cartesian coordinates (x, y) by

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

For concreteness, we assume (and will *always* assume) the angle θ is measured in radians, so the range $\theta \in [0, 2\pi]$ describes a full rotation. In polar coordinates, S^1 is the subset $\{r = 1\} \subset \mathbb{R}^2$, thus one of the coordinates becomes irrelevant, and having one coordinate left makes S^1 a one-dimensional object.

The above discussion of polar coordinates glossed over an important point: one cannot simultaneously describe *every* point in S^1 via a unique value of the angular coordinate $\theta \in \mathbb{R}$, at least not if we want the values of θ to be unambiguously defined and continuously dependent on the points that they describe. One could e.g. require θ to take values only in a half-open interval like $[0, 2\pi)$ or $(-\pi, \pi]$: this creates a one-to-one correspondence between points on S^1 and values of the coordinate, but the function one defines in this way from S^1 to $[0, 2\pi)$ or $(-\pi, \pi]$ has a jump discontinuity at the point where the coordinate reaches either end of the allowed interval. If you want to avoid such discontinuities, then the only option is to give up on the notion of describing *all* of S^1 in a single coordinate system, and instead use multiple coordinate systems defined on different subsets. For instance, we could define two subsets of the circle by

$$\mathcal{U} := S^1 \backslash \{(1, 0)\}, \qquad \mathcal{V} := S^1 \backslash \{(-1, 0)\},$$

and associate to these two subsets two potentially different angular coordinates θ and ϕ respectively, each taking values in an appropriate open interval, thus defining *continuous* functions

$$\theta : \mathcal{U} \to (0, 2\pi), \qquad \phi : \mathcal{V} \to (-\pi, \pi).$$

Since $S^1 = \mathcal{U} \cup \mathcal{V}$, these two coordinate systems together can be used to describe every point in S^1. Moreover, there is a large region on which both coordinates θ and ϕ are defined: it consists of the two semi-circles $S^1_+ := \{(x,y) \in S^1 \mid y > 0\}$ and $S^1_- := \{(x,y) \in S^1 \mid y < 0\}$, and on each of these one can easily derive a relationship between θ and ϕ, namely

$$\phi = \begin{cases} \theta & \text{on } S^1_+, \\ \theta - 2\pi & \text{on } S^1_-. \end{cases} \tag{1.4}$$

The pairs $(\mathcal{U}, \theta)$ and $(\mathcal{V}, \phi)$ are our first examples of what we will call *charts* on the 1-dimensional manifold S^1, and together they form a *smooth atlas* that determines a *smooth structure* on S^1. Let us now begin giving precise definitions to these terms.

In the following, assume M is a set, and $n \geqslant 0$ is an integer. For the sake of intuition, you may picture M as a surface (in which case $n = 2$), and picture the subsets $\mathcal{U}, \mathcal{V} \subset M$ as open subsets of that surface.[8] Recall that a continuous map defined on an open subset of Euclidean space is called **smooth** (*glatt*) if it admits derivatives of all orders.

Definition 1.4. An n-dimensional **chart** (*Karte*)[9] $(\mathcal{U}, x)$ on M consists of a subset $\mathcal{U} \subset M$ and an injective map $x : \mathcal{U} \hookrightarrow \mathbb{R}^n$ whose image $x(\mathcal{U}) \subset \mathbb{R}^n$ is an open set.

Any two charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ determine a pair of **transition maps** (*Kartenübergänge*)

$$\begin{aligned} \mathbb{R}^n \supset x(\mathcal{U} \cap \mathcal{V}) &\xrightarrow{y \circ x^{-1}} y(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{R}^n, \\ \mathbb{R}^n \supset y(\mathcal{U} \cap \mathcal{V}) &\xrightarrow{x \circ y^{-1}} x(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{R}^n, \end{aligned} \tag{1.5}$$

which are inverse to each other, and are thus bijections between subsets of $\mathbb{R}^n$. We say that the two charts are C^k**-compatible** (*verträglich*) for some $k \in \mathbb{N} \cup \{0, \infty\}$ if the sets $x(\mathcal{U} \cap \mathcal{V})$ and $y(\mathcal{U} \cap \mathcal{V})$ are both open and the transition maps $y \circ x^{-1}$ and $x \circ y^{-1}$ are both of class C^k. If $k = \infty$, we say the charts are **smoothly compatible** (*glatt verträglich*).

A picture of what a pair of overlapping charts on a surface might look like is shown in Figure 4. An individual chart $(\mathcal{U}, x)$ should be understood as defining a *coordinate system* for describing all points in the subset $\mathcal{U} \subset M$, where the individual **coordinates** (*Koordinaten*) are the n real-valued functions

$$x^1, \ldots, x^n : \mathcal{U} \to \mathbb{R}$$

defined as the component functions of the map $x = (x^1, \ldots, x^n) : \mathcal{U} \to \mathbb{R}^n$. Note that in Definition 1.4, it is permissible for the domains $\mathcal{U}$ and $\mathcal{V}$ of the two charts to be disjoint, in which case the transition maps $y \circ x^{-1}$ and $x \circ y^{-1}$ are both just the trivial map from the empty set to itself. But if $\mathcal{U} \cap \mathcal{V} \neq \varnothing$, then the transition map

$$\begin{array}{ccc} & \mathcal{U} \cap \mathcal{V} & \\ {}^{x}\swarrow & & \searrow^{y} \\ \mathbb{R}^n \overset{\text{open}}{\supset} x(\mathcal{U} \cap \mathcal{V}) & \xrightarrow{\quad y \circ x^{-1} \quad} & y(\mathcal{U} \cap \mathcal{V}) \overset{\text{open}}{\subset} \mathbb{R}^n \end{array}$$

[8]Saying the word "open" presumes that M has some structure beyond merely being an arbitrary set, e.g. it could be a subset of some Euclidean space $\mathbb{R}^n$, or more generally, a metric or topological space. We will address this point properly in the next lecture, but since we have not addressed it yet, Definition 1.4 refers to $\mathcal{U}$ and $\mathcal{V}$ simply as "subsets" of M, without saying they are open. In practice, they always will be.

[9]A word of caution for German speakers: the mathematical word *Abbildung* (as in "eine injektive Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$") can be translated into English as either "map" or "mapping", but do not be tempted to translate "map" into mathematical German as *Karte*. In mathematical English, a "chart" and a "map" are not exactly the same thing.

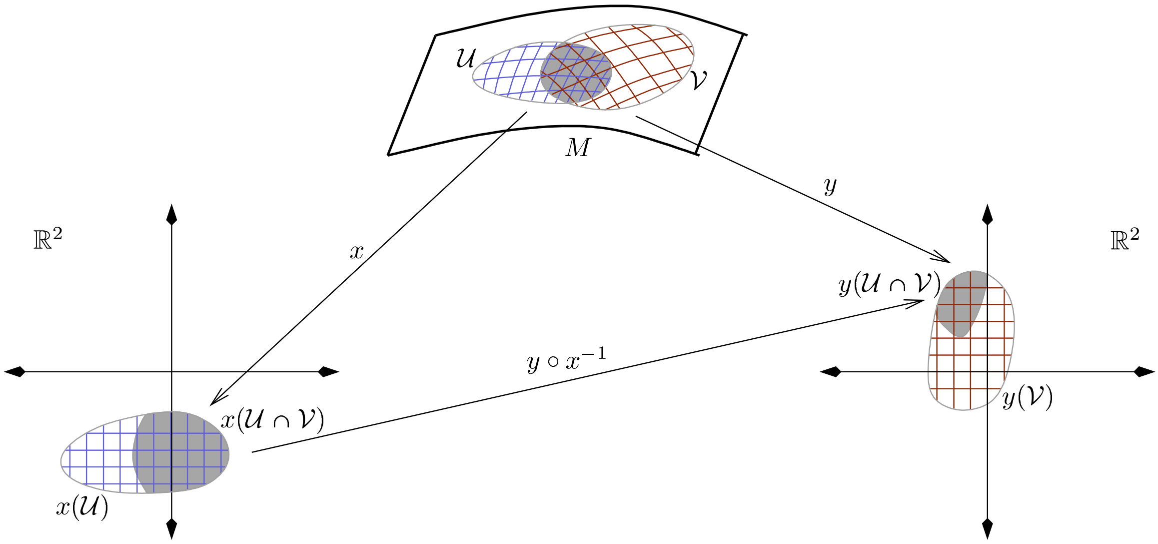


FIGURE 4. Two charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ on a surface M, with an associated transition map $y \circ x^{-1}$ defining a bijection between two open sets (the shaded regions) in $\mathbb{R}^2$.

defines a coordinate transformation, e.g. for any point $p \in \mathcal{U} \cap \mathcal{V}$, $y \circ x^{-1}$ sends the vector $(x^1(p), \ldots, x^n(p)) \in \mathbb{R}^n$ that represents p in "x-coordinates" to the vector that represents the same point in "y-coordinates", namely $(y^1(p), \ldots, y^n(p)) \in \mathbb{R}^n$. It is often convenient in this situation to write the y-coordinates on the overlap region as functions of the x-coordinates, i.e. if we identify each point in $\mathcal{U} \cap \mathcal{V}$ with the vector in $\mathbb{R}^n$ determined by its x-coordinates, then the y-coordinates can be viewed as functions of n variables, which are naturally labelled $x^1, \ldots, x^n$, producing a transformation

$$(x^1, \ldots, x^n) \mapsto (y^1(x^1, \ldots, x^n), \ldots, (y^n(x^1, \ldots, x^n)). \tag{1.6}$$

This is a slight abuse of notation, because in this expression, the variables $x^1, \ldots, x^n$ are no longer interpreted as real-valued functions on $\mathcal{U} \subset M$, but simply as the usual Cartesian coordinates on the open subset $x(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{R}^n$. With this understood, (1.6) is just another expression for the transition map $y \circ x^{-1}$, and the inverse transition map $x \circ y^{-1}$ can similarly be written as

$$(y^1, \ldots, y^n) \mapsto (x^1(y^1, \ldots, y^n), \ldots, (x^n(y^1, \ldots, y^n)), \tag{1.7}$$

with the variables $y^1, \ldots, y^n$ now understood to represent Cartesian coordinates on $y(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{R}^n$. If the two charts are C^k-compatible, then both of the transformations in (1.6) and (1.7) are of class C^k. If $k \geqslant 1$, then since the two transformations are inverse to each other, it follows that the n-by-n matrix with entries

$$\frac{\partial y^i}{\partial x^j}(x^1, \ldots, x^n), \qquad i, j \in \{1, \ldots, n\}$$

is invertible for every $(x^1, \ldots, x^n) \in x(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{R}^n$.

REMARK 1.5. You may have been accustomed to using *subscripts* $x_1, \ldots, x_n$ for coordinates on $\mathbb{R}^n$ in your studies up to this point, and will thus wonder why I am instead using superscripts in all the expressions above. This is not an arbitrary choice—it is a convention that is widespread in differential geometry, and especially popular among physicists, and we will try to use it consistently throughout this course. Subscripts will at some point also appear, but they will have a different meaning.

EXAMPLE 1.6. In the discussion of the unit circle S^1 above, we defined two charts $(\mathcal{U}, \theta)$ and $(\mathcal{V}, \phi)$, with images $\theta(\mathcal{U}) = (0, 2\pi) \subset \mathbb{R}$ and $\phi(\mathcal{V}) = (-\pi, \pi) \subset \mathbb{R}$. The overlap region $\mathcal{U} \cap \mathcal{V}$ of these two charts is the union of two disjoint open sets that we denoted by S^1_+ and S^1_-, the upper and lower semicircle (disjoint from the x-axis). The transition map $\phi \circ \theta^{-1} : \theta(S^1_+ \cup S^1_-) \to \phi(S^1_+ \cup S^1_-)$ is then found by writing ϕ as a function of θ as in (1.4), which gives

$$\phi(\theta) = \begin{cases} \theta & \text{for } 0 < \theta < \pi, \\ \theta - 2\pi & \text{for } \pi < \theta < 2\pi. \end{cases}$$

Observe that while this map appears at first glance to have a jump discontinuity, its actual domain is $\theta(S^1_+ \cup S^1_-) = (0, \pi) \cup (\pi, 2\pi)$, i.e. it excludes the point π at which the discontinuity would occur. As a result, this transition map is smooth, and so is its inverse; the two charts $(\mathcal{U}, \theta)$ and $(\mathcal{V}, \phi)$ are therefore smoothly compatible.

EXERCISE 1.7. The standard **spherical coordinates** (*Kugelkoordinaten*) on $\mathbb{R}^3$ are defined via the transformation

$$(r, \theta, \phi) \mapsto (x, y, z), \qquad \begin{cases} x & := r \cos\theta \cos\phi, \\ y & := r \sin\theta \cos\phi, \\ z & := r \sin\phi, \end{cases} \tag{1.8}$$

where θ plays the role of an angle in the xy-plane, and $\phi \in [-\pi/2, \pi/2]$ is the angle between the vector $(x, y, z) \in \mathbb{R}^3$ and the xy-plane.[10] Restricting to $r = 1$, the other two coordinates (θ, ϕ) can be used to describe points on the unit sphere $S^2 \subset \mathbb{R}^3$, though there are choices to be made since θ is only defined up to multiples of 2π (and it is not defined at all at the north and south poles $p_\pm := (0, 0, \pm 1) \in S^2$, where $\phi = \pm\pi/2$.)

(a) Find two subsets $\mathcal{U}_1, \mathcal{U}_2 \subset S^2$ with $\mathcal{U}_1 \cup \mathcal{U}_2 = S^2 \backslash \{p_+, p_-\}$ such that for $i = 1, 2$, there are 2-dimensional charts of the form $(\mathcal{U}_i, \alpha_i)$ with $\alpha_i = (\theta_i, \phi_i)$, where the coordinate functions $\theta_i, \phi_i : \mathcal{U}_i \to \mathbb{R}$ are continuous and satisfy the spherical coordinate relations (1.8), and have images $\alpha_1(\mathcal{U}_1) = (0, 2\pi) \times (-\pi/2, \pi/2) \subset \mathbb{R}^2$ and $\alpha_2(\mathcal{U}_2) = (-\pi, \pi) \times (-\pi/2, \pi/2) \subset \mathbb{R}^2$.

(b) One cannot use spherical coordinates to construct a chart on S^2 that contains either of the poles $p_\pm = (0, 0, \pm 1)$. Can you think of another way to construct charts on open subsets of S^2 that contain these two points?
Hint: On any sufficiently small neighborhood of p_+ or p_- in S^2, every point has its z-coordinate determined by the x and y-coordinates.

(c) Now that you've constructed charts that cover every point on S^2, write down the associated transition maps and show that your charts are all smoothly compatible with each other.

2. Smooth manifolds

In this lecture we give the definition of the term *smooth manifold* and look at a few more examples.

---

[10] Achtung: there are various conventions for spherical coordinates in use. I'm told that this is the standard convention learned by mathematics students in Germany. I learned a different convention as a physics student in the U.S.: $x = r \cos\phi \sin\theta$, $y = r \sin\phi \sin\theta$, $z = r \cos\theta$. Here ϕ plays the role of the angle in the xy-plane, and $\theta \in [0, \pi]$ is the angle between $(x, y, z) \in \mathbb{R}^3$ and the positive z-axis.

2.1. Atlases and smooth structures. We concluded Lecture 1 by defining the notion of a *chart* on a set M, and C^k-compatibility between two charts. A chart $(\mathcal{U}, x)$ should be interpreted as a "local" coordinate system, which can be used to label points in the subset $\mathcal{U} \subset M$. We saw in the example of the circle S^1 that while one cannot apparently describe *all* points in S^1 via a single chart, it was easy to find two smoothly compatible charts such that every point is in at least one or the other. Exercise 1.7 similarly outlines how to cover S^2 with four charts using spherical coordinates. These were the first examples of the following general concept.

DEFINITION 2.1. An **atlas of class** C^k for the set M (or **smooth atlas** in the case $k = \infty$) is a collection of charts $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ that are all C^k-compatible with each other, such that $\bigcup_{\alpha \in I} \mathcal{U}_\alpha = M$.[11]

In first-year analysis, you learned what it means for a real-valued function on an open subset of $\mathbb{R}^n$ to be differentiable; it was important in that definition that the domain of the function should be *open*, as differentiation at a point p involves limits that are not well defined unless f itself is defined on some ball around p. In differential geometry, we would also like to be able to differentiate functions

$$f : M \to \mathbb{R}$$

defined on a manifold M, such as the circle S^1 or sphere S^2. This is a nontrivial problem, even in simple examples such as S^n that are given as subsets of Euclidean space, since they are not generally *open* subsets. But if M is a set equipped with an atlas, then M is covered by subsets that have coordinate systems, so for each chart $(\mathcal{U}, x)$ we can write down f "in local coordinates", meaning we identify each point $p \in \mathcal{U}$ with its coordinate vector $(x^1(p), \dots, x^n(p)) \in \mathbb{R}^n$, so that $f|_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}$ becomes a function of n real variables

$$(x^1, \dots, x^n) \mapsto f(x^1, \dots, x^n), \tag{2.1}$$

with $x^1, \dots, x^n$ interpreted as the standard Cartesian coordinates on the open set $x(\mathcal{U}) \subset \mathbb{R}^n$. This is another slight abuse of notation, similar to the coordinate expressions for transition maps described in (1.6) and (1.7); in fact, the function that is literally described in (2.1) is not $f : M \to \mathbb{R}$ but rather

$$x(\mathcal{U}) \xrightarrow{f \circ x^{-1}} \mathbb{R}.$$

It now seems natural to say that f is differentiable at $p \in \mathcal{U} \subset M$ if and only if its coordinate expression $f \circ x^{-1}$ is differentiable (in the sense of first-year analysis) at the corresponding point $x(p) \in x(\mathcal{U}) \subset \mathbb{R}^n$. For this to be a reasonable definition, we need to know that it does not depend on the *choice* of the chart $(\mathcal{U}, x)$, as our atlas may indeed contain multiple distinct charts that contain the point p. This issue is precisely what the compatibility condition in Definition 1.4 was designed to settle:

LEMMA 2.2. *Suppose $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ are two C^k-compatible charts on M, and $f : M \to \mathbb{R}$ is a function. Then for each nonnegative integer $r \leqslant k$, the function $x(\mathcal{U} \cap \mathcal{V}) \xrightarrow{f \circ x^{-1}} \mathbb{R}$ is of class C^r if and only if the function $y(\mathcal{U} \cap \mathcal{V}) \xrightarrow{f \circ y^{-1}} \mathbb{R}$ is of class C^r.*

PROOF. The statement follows from the chain rule, since $f \circ y^{-1} = (f \circ x^{-1}) \circ (x \circ y^{-1})$ and $f \circ x^{-1} = (f \circ y^{-1}) \circ (y \circ x^{-1})$. ☐

DEFINITION 2.3. For a set M with an atlas $\mathcal{A}$ of class C^k and $r \in \mathbb{N} \cup \{0, \infty\}$ with $r \leqslant k$, a function $f : M \to \mathbb{R}$ is said to be **of class** C^r if and only if the function $x(\mathcal{U}) \xrightarrow{f \circ x^{-1}} \mathbb{R}$ is of class C^r for every chart $(\mathcal{U}, x) \in \mathcal{A}$.

[11] In this definition, I may be any set, finite, countable or uncountable. We refer to it as an **index set** since it is only used for labelling purposes and is otherwise unimportant in itself.

EXERCISE 2.4. Convince yourself that Lemma 2.2 becomes false in general if one allows $r > k$. (See also Example 2.7 below for a concrete special case.) This has the following consequence: if we want to define what it means for a function on a manifold to be of class C^k, then we need to have an atlas of class C^k *or better* to test it with. In particular, the notion of smooth functions on M cannot be defined unless M is equipped with a *smooth* atlas.

The examples of smooth atlases we saw in Lecture 1 on S^1 and S^2 were finite, and this will turn out to be a general pattern: we will see that almost all manifolds we are interested in admit finite atlases, though it is not often important to know this. On the other hand, a general atlas can be uncountably infinite, and one can always enlarge a finite atlas $\{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ in trivial ways, e.g. by choosing subsets $\mathcal{U}'_\alpha \subset \mathcal{U}_\alpha$ for which $x_\alpha(\mathcal{U}'_\alpha) \subset \mathbb{R}^n$ is open and adding in the restricted charts $(\mathcal{U}'_\alpha, x_\alpha|_{\mathcal{U}'_\alpha})$, which are obviously still compatible with all the others. We say that an atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ of class C^k is **maximal** if it cannot be enlarged any further without sacrificing compatibility, i.e. every chart that is C^k-compatible with all of the charts in $\mathcal{A}$ already belongs to $\mathcal{A}$.

LEMMA 2.5. *Given an atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ of class C^k on M, let $\mathcal{A}'$ denote the collection of all charts on M that are C^k-compatible with all the charts in $\mathcal{A}$. Then $\mathcal{A}'$ is a maximal atlas of class C^k, and it is the only one containing $\mathcal{A}$.*

PROOF. To show that $\mathcal{A}'$ is an atlas, we need to show that any two charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ that are C^k-compatible with every $(\mathcal{U}_\alpha, x_\alpha)$ are also C^k-compatible with each other. Given a point $p \in \mathcal{U} \cap \mathcal{V}$, pick $\alpha \in I$ so that $p \in \mathcal{U}_\alpha$. The set $x(\mathcal{U} \cap \mathcal{V} \cap \mathcal{U}_\alpha) \subset \mathbb{R}^n$ is then the intersection of the two open sets $x(\mathcal{U} \cap \mathcal{U}_\alpha)$ and $x(\mathcal{V} \cap \mathcal{U}_\alpha)$ and is thus an open neighborhood of $x(p)$, so on this neighborhood, the transition map $y \circ x^{-1}$ can then be written as

$$y \circ x^{-1} = (y \circ x_\alpha^{-1}) \circ (x_\alpha \circ x^{-1}),$$

which is a composition of two C^k-maps and is therefore of class C^k on the neighborhood of $x(p)$ in question. This trick works (possibly with different choices of α) for any point $p \in \mathcal{U} \cap \mathcal{V}$, and it also works for the inverse transition map $x \circ y^{-1}$, thus it implies that both of the transition maps relating x and y are everywhere of class C^k, and $\mathcal{A}'$ is therefore an atlas. It clearly also contains $\mathcal{A}$, and it is maximal, since any chart compatible with every chart in $\mathcal{A}'$ is also compatible with every chart in $\mathcal{A}$, and thus belongs to $\mathcal{A}'$ by definition. Finally, if $\mathcal{A}''$ is any other atlas containing $\mathcal{A}$, then every chart in $\mathcal{A}''$ is compatible with every chart in $\mathcal{A} \subset \mathcal{A}''$ and therefore belongs to $\mathcal{A}'$ by definition, proving $\mathcal{A}'' \subset \mathcal{A}'$. If $\mathcal{A}''$ is also maximal, it follows that $\mathcal{A}'' = \mathcal{A}'$. □

DEFINITION 2.6. For $k \in \mathbb{N} \cup \{\infty\}$, a **$C^k$-structure** (*$C^k$-Struktur*) or **differentiable structure of class C^k** (*differenzierbare Struktur von der Klasse C^k*) on a set M is a maximal atlas $\mathcal{A}$ of class C^k on M. In the case $k = \infty$, we also call this a **smooth structure** (*glatte Struktur*) on M. If M has been endowed with a C^k-structure $\mathcal{A}$, then a chart $(\mathcal{U}, x)$ on M will be referred to as a **C^k-chart** (or a **smooth chart** in the case $k = \infty$) if it belongs to the maximal atlas $\mathcal{A}$.

The maximality condition in Definition 2.6 is convenient for bookkeeping purposes (see Remark 2.8 below), but Lemma 2.5 shows that it is not a meaningful restriction. In practice, one typically specifies a smooth structure by first describing the smallest atlas one is able to construct, and then replacing it with its unique maximal extension. We will usually carry out the latter step without even mentioning it.

EXAMPLE 2.7. The following defines an atlas of class C^0 but not C^1 on $\mathbb{R}$: consider two charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ with

$$\mathcal{U} := (-\infty, 1), \qquad x(t) := t,$$
$$\mathcal{V} := (-1, \infty), \qquad y(t) := t^3.$$

The resulting transition maps both send $(-1,1) \to (-1,1)$ and are given by

$$y(x) = x^3, \qquad x(y) = \sqrt[3]{y},$$

so both are continuous, but $x \circ y^{-1}$ is not differentiable. This has the consequence that functions $\mathbb{R} \to \mathbb{R}$ that look differentiable in the x-coordinate might not look differentiable in the y-coordinate. An easy example is the identity map $f(t) = t$, which looks like $f(x) = x$ and is thus smooth in the x-coordinate, but its expression in the y-coordinate is $f(y) = \sqrt[3]{y}$, which fails to be differentiable at the point $0 \in y(\mathcal{V}) = (-1, \infty)$.

Note that if we enlarge both $\mathcal{U}$ and $\mathcal{V}$ to $\mathbb{R}$, then while the two charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ together do not determine any smooth structure on $\mathbb{R}$, each of these charts individually forms a smooth atlas—an atlas with only one chart is always smooth since it has no nontrivial transition maps whose differentiability would need to be checked. Each therefore determines a smooth structure via Lemma 2.5, and in this way, one obtains two *different* smooth structures on $\mathbb{R}$.

Remark 2.8. The advantage of requiring maximality in Definition 2.6 is the following: if $\mathcal{A}$ and $\mathcal{A}'$ are two atlases on M for which every chart in $\mathcal{A}$ is compatible with every chart in $\mathcal{A}'$, then the two notions of differentiability for functions on M defined via these two atlases will be the same, and we would therefore prefer to think of them is defining the *same* smooth structure, even if they are different atlases, strictly speaking. In this scenario, it is easy to check that both atlases do in fact have the same maximal extension.

2.2. Some topological notions. With the concept of a smooth atlas in hand, a reasonable guess for the "right" definition of a smooth manifold would be that it is any set endowed with the additional structure of a smooth atlas. In practice, however, doing anything interesting with manifolds requires imposing one or two further restrictions on what is allowed to be a manifold and what is not.

I do not want to assume previous knowledge of topology in this course, but a few basic notions of the subject now need to be discussed before we can give the precise definition of a manifold. Most of them will play a negligible role in this course, and in fact, the intuition you already have about metric spaces is fully sufficient for understanding the definition of a manifold (cf. Remark 2.20 below)—nonetheless, you will not be able to understand *why* that definition is what it is unless we first discuss the alternatives.

Since you have seen metric spaces before, you know how to define fundamental notions such as **continuity** (*Stetigkeit*), **convergence** of a sequence to a point (*Konvergenz einer Folge gegen einen Punkt*) and **closed** sets (*abgeschlossene Teilmengen*) in metric spaces. You will also have seen important concepts such as that of a **neighborhood** (*Umgebung*) of a point $x \in X$, meaning any subset $\mathcal{U} \subset X$ that contains an open subset containing x, and probably also a **homeomorphism** (*Homöomorphismus*), which is a continuous bijection whose inverse is also continuous. One detail you may or may not already be aware of is that all of these notions can be defined without any explicit reference to a metric, so long as one knows what an "open set" is. In particular:

Proposition 2.9 (first-year analysis). *Assume X and Y are metric spaces.*

(1) *A sequence $x_n \in X$ converges to a point $x \in X$ if and only if for every neighborhood $\mathcal{U} \subset X$ of x, $x_n \in \mathcal{U}$ for all sufficiently large n.*
(2) *A subset $\mathcal{U} \subset X$ is closed if and only if its complement $X \backslash \mathcal{U} \subset X$ is open.*
(3) *A map $f : X \to Y$ is continuous if and only if for every open subset $\mathcal{U} \subset Y$, $f^{-1}(\mathcal{U}) := \{x \in X \mid f(x) \in \mathcal{U}\}$ is an open subset of X.*
(4) *A bijective map $f : X \to Y$ is a homeomorphism if and only if it defines a bijective correspondence between the open subsets of X and the open subsets of Y, i.e. for all subsets $\mathcal{U} \subset X$, $\mathcal{U}$ is open if and only if $f(\mathcal{U}) \subset Y$ is open.*

□

Exercise 2.10. If you do not already find Proposition 2.9 obvious, prove it.

Topology begins with the observation that it is sometimes convenient to define what an open set is without the aid of a metric. For this idea to be useful, we just need open sets to satisfy a few properties that are already familiar from the theory of metric spaces:

Definition 2.11. A **topology** (*Topologie*) on a set X is a collection $\mathcal{T}$ of subsets of X satisfying the following axioms:

(i) $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$;
(ii) For every subcollection $I \subset \mathcal{T}$, $\bigcup_{\mathcal{U} \in I} \mathcal{U} \in \mathcal{T}$;
(iii) For every pair $\mathcal{U}_1, \mathcal{U}_2 \in \mathcal{T}$, $\mathcal{U}_1 \cap \mathcal{U}_2 \in \mathcal{T}$.

The pair $(X, \mathcal{T})$ is then called a **topological space** (*topologischer Raum*), and we call the sets $\mathcal{U} \in \mathcal{T}$ the **open** subsets (*offene Teilmengen*) in $(X, \mathcal{T})$.

We will usually not give an actual label to the topology when discussing a topological space, so e.g. instead of talking about $(X, \mathcal{T})$, we will talk about "the topological space X" with the understanding that a subset $\mathcal{U} \subset X$ is called "open" if and only if it belongs to the topology that has been specified on X. For topological spaces X and Y, one now takes the statements in Proposition 2.9 as *definitions* of the notions of convergence, closed subsets, continuity and homeomorphisms.

We call a topological space X **metrizable** (*metrisierbar*) if it admits a metric for which the given topology of X consists of all sets that are unions of open balls, i.e. the metrizable spaces are the topological spaces that you already saw (but without using the word "topology") when you studied metric spaces. Two things about this notion are important to understand:

(1) If X is metrizable, then the metric that defines its topology is typically far from being unique. For example, $d(x, y) := c|x - y|$ for any constant $c > 0$ defines a "nonstandard" metric on $\mathbb{R}$ that nonetheless induces the same topology as the standard one.
(2) Many topological spaces are not metrizable, and they can easily have properties that are counterintuitive. (We will see an example in a moment.)

We saw in §2.1 that an atlas of class C^k on a set M determines a natural way to define what it means for a function $f : M \to \mathbb{R}$ to be of class C^r for any $r \leqslant k$. This holds in particular for $r = 0$, so that continuity of functions can be defined in a certain sense, even though we never explicitly endowed M with a topology. But actually, we did, we just didn't notice:

Proposition 2.12. *Given an atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ of class C^0 on a set M, there exists a unique topology on M such that the sets $\mathcal{U}_\alpha \subset M$ are all open and the maps x_α are all homeomorphisms onto their images.*[12] *Moreover, for every other chart $(\mathcal{U}, x)$ that is C^0-compatible with the charts in $\mathcal{A}$, $\mathcal{U} \subset M$ is also open and x is also a homeomorphism onto its image.*

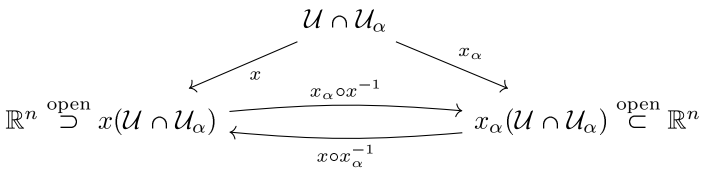
Proof. Suppose M carries a topology with the properties described, and $\mathcal{O} \subset M$ is an open subset. Then each of the sets $\mathcal{O}_\alpha := \mathcal{O} \cap \mathcal{U}_\alpha$ is open, and $\mathcal{O} = \bigcup_{\alpha \in I} \mathcal{O}_\alpha$. Since each x_α is a homeomorphism onto its image in $\mathbb{R}^n$, $x_\alpha(\mathcal{O}_\alpha)$ is then also an open subset of $\mathbb{R}^n$. Conversely, if $\mathcal{O} \subset M$ is *any* subset such the sets $\Omega_\alpha := x_\alpha(\mathcal{O} \cap \mathcal{U}_\alpha) \subset \mathbb{R}^n$ are all open, then each $\mathcal{O}_\alpha := \mathcal{O} \cap \mathcal{U}_\alpha = x_\alpha^{-1}(\Omega_\alpha) \subset M$ must also be open since x_α is a homeomorphism, and therefore so is the union $\mathcal{O} = \bigcup_{\alpha \in I} \mathcal{O}_\alpha$. This proves that a topology with the stated properties is unique: if it exists,

[12] Recall that $x_\alpha(\mathcal{U}_\alpha)$ is an open subset of a Euclidean space $\mathbb{R}^n$, so it is understood in this statement to carry the obvious topology that it inherits from the Euclidean metric on $\mathbb{R}^n$.

then it is precisely the collection of all subsets $\mathcal{O} \subset M$ such that $x_\alpha(\mathcal{O} \cap \mathcal{U}_\alpha) \subset \mathbb{R}^n$ is open for every $\alpha \in I$.

To prove existence, one now has to prove that the collection of subsets of M described above satisfies the axioms of a topology, i.e. it contains M and $\varnothing$ and is closed under arbitrary unions and finite intersections. This is a straightforward exercise.

Finally, let us fix the topology on M described above and suppose $(\mathcal{U}, x)$ is another chart that is C^0-compatible with $(\mathcal{U}_\alpha, x_\alpha)$ for every $\alpha \in I$. We need to show that $\mathcal{U} \subset M$ is open and $x : \mathcal{U} \to \mathbb{R}^n$ is a homeomorphism onto its image, which is equivalent to showing that for subsets $\mathcal{O} \subset \mathcal{U}$, $\mathcal{O}$ is open in M if and only if $x(\mathcal{O})$ is open in $\mathbb{R}^n$. For this, we make use of the transition maps relating $(\mathcal{U}, x)$ and $(\mathcal{U}_\alpha, x_\alpha)$ for an arbitrary choice of $\alpha \in I$:



By the assumption of C^0-compatibility, the two maps in the bottom row of this diagram are both continuous, and since they are inverse to each other, they are homeomorphisms, meaning they define a bijection between the open subsets of $x(\mathcal{U} \cap \mathcal{U}_\alpha)$ and $x_\alpha(\mathcal{U} \cap \mathcal{U}_\alpha)$. Now suppose $\mathcal{O} \subset M$ is open, which means $x_\alpha(\mathcal{O} \cap \mathcal{U}_\alpha) \subset x_\alpha(\mathcal{U} \cap \mathcal{U}_\alpha) \subset \mathbb{R}^n$ is open for every α. Feeding this set into the homeomorphism $x \circ x_\alpha^{-1}$ gives $x(\mathcal{O} \cap \mathcal{U}_\alpha)$, proving that the latter is an open set, and therefore so is $x(\mathcal{O}) = \bigcup_{\alpha \in I} x(\mathcal{O} \cap \mathcal{U}_\alpha)$. Conversely, if $\mathcal{O} \subset M$ is an arbitrary subset such that $x(\mathcal{O})$ is open, then for every $\alpha \in I$, $x(\mathcal{O} \cap \mathcal{U}_\alpha)$ is the intersection of two open sets $x(\mathcal{O})$ and $x(\mathcal{U} \cap \mathcal{U}_\alpha)$, and is thus also open. Feeding it into $x_\alpha \circ x^{-1}$ then shows that $x_\alpha(\mathcal{O} \cap \mathcal{U}_\alpha)$ is also open, proving that $\mathcal{O} \subset M$ is open. □

Whenever we discuss a set M with an atlas $\mathcal{A}$ from now on, we will assume that M is endowed with the topology described in Proposition 2.12.

Remark 2.13. Notice that according to the last statement in Proposition 2.12, the topologies induced on M by $\mathcal{A}$ or any extension of $\mathcal{A}$ to a larger (e.g. maximal) atlas are the same.

Remark 2.14. It is rarely actually necessary to apply Proposition 2.12 for defining a topology on a manifold. The much more common situation is that our manifold M comes equipped with some natural topology that is clear from the context (e.g. because M is a subset or quotient of $\mathbb{R}^n$ or some other manifold that we already understand), and when specifying an atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ for M, we just need to check that the topology determined by the atlas is the same as the natural topology. In other words, we need to check that the sets $\mathcal{U}_\alpha$ are open and the maps $x_\alpha : \mathcal{U}_\alpha \to x_\alpha(\mathcal{U}_\alpha) \subset \mathbb{R}^n$ are all homeomorphisms with respect to the natural topology. In most situations, this will be obvious.

Exercise 2.15. We now have two ways of defining what it means for a function $f : M \to \mathbb{R}$ to be continuous: one is the case $k = 0$ of Definition 2.3, in terms of the atlas $\mathcal{A}$, and the other is the standard notion of continuity in topological spaces, using the topology determined by $\mathcal{A}$ according to Proposition 2.12. Convince yourself that these two definitions are equivalent.

Since the atlas identifies small neighborhoods in M with neighborhoods in Euclidean space, and the topology of Euclidean space is pleasantly familiar to us, one might intuitively expect the topology induced on M by $\mathcal{A}$ to have similarly pleasant properties. The next example shows that this intuition is wrong.

EXAMPLE 2.16. Define an equivalence relation $\sim$ on the set $\widetilde{M} := \mathbb{R} \times \{0,1\}$ such that every element is equivalent to itself and $(t,0) \sim (t,1)$ for all $t \in \mathbb{R}\backslash\{0\}$, but not for $t = 0$. Let

$$M := \widetilde{M}/\sim$$

denote the set of equivalence classes. We can think of M intuitively as a "real line with two zeroes", because it mostly looks just the same as $\mathbb{R}$ (each number $t \neq 0$ corresponding to the equivalence class of $(t,0)$ and $(t,1)$), but $t = 0$ is an exception, where there really are *two* distinct points $[(0,0)]$ and $[(0,1)]$ in M. The following pair of 1-dimensional charts define a smooth atlas on M: let

$$\mathcal{U}_\alpha := \left\{[(t,0)] \in M \mid t \in \mathbb{R}\right\}, \qquad \mathcal{U}_\beta := \left\{[(t,1)] \in M \mid t \in \mathbb{R}\right\},$$

and define both $x_\alpha : \mathcal{U}_\alpha \to \mathbb{R}$ and $x_\beta : \mathcal{U}_\beta \to \mathbb{R}$ by $[(t,k)] \mapsto t$ for $k = 0,1$. The transition maps relating these two charts are both the identity map on $\mathbb{R}\backslash\{0\}$, thus the charts are smoothly compatible, and clearly $M = \mathcal{U}_\alpha \cup \mathcal{U}_\beta$.

Now consider the sequence

$$p_j := [(1/j, 0)] \in M.$$

Does it converge? We need to think for a moment about what convergence means in the topology induced by an atlas: if $p \in \mathcal{U}_\alpha$, then since x_α is a homeomorphism onto its image, p_j will converge to p if and only if $x_\alpha(p_j)$ converges to $x_\alpha(p)$ in $\mathbb{R}$, and a moment's thought reveals that that condition holds for $p := [(0,0)]$. However, if we use the *other* chart x_β, then since $(1/j, 0) \sim (1/j, 1)$ for every j, the same condition also holds for the point $p' := [(0,1)] \in \mathcal{U}_\beta$, and we have thus found two *distinct* points $p \neq p'$ such that $p_j \to p$ and $p_j \to p'$.

This seems like a contradiction if you have not seen any topology before, but it is not: it merely shows that M is a much stranger topological space than our intuition about metric spaces had led us to expect. In fact, the points p and p' have the peculiar property that every neighborhood of p intersects every neighborhood of p', so even though they are distinct points, the topology of M does not "separate" them; the technical term for this is that the topology of M is not **Hausdorff**.[13]

We do not want our notion of manifolds to include pathological examples in which a sequence can converge to two distinct points at once. Among other issues, it would clearly be impossible to define a metric compatible with that notion of convergence, as the triangle inequality ensures that limits of sequences are unique in metric spaces. Since the notion of distance on manifolds is one of the main things we plan to study when we get further into this subject, we would like to have a guarantee that every manifold *admits* a metric that is compatible with its natural topology, i.e. we will insist that all manifolds be metrizable. This condition will turn out to have many advantages beyond the study of distance, though we will rarely need to make explicit use of it: it will only become important when we discuss the construction of global geometric structures (such as Riemannian metrics) via partitions of unity.

Although it will play no significant role in this course, we need one more topological notion in order to understand the main definition: a topological space is called **separable** (*separabel*) if it contains a countable dense subset. Euclidean spaces, for example, are separable, because $\mathbb{Q}^n \subset \mathbb{R}^n$ is a countable dense subset. Every space of interest in this course will be separable, and one can often use the result of the following exercise to prove it.

EXERCISE 2.17. Show that every subset of a separable metric space (X,d) is also a separable metric space.
Hint: Given a countable dense subset $E \subset X$ and another subset $Y \subset X$, show first that every open set in X is a union of open balls of the form $B_r(x) := \left\{y \in X \mid d(y,x) < r\right\}$ for $x \in E$ and

[13]Or, as my topology professor in grad school once put it, the points p and p' are not "housed off" from each other. The proper delivery of this joke requires a Brooklyn accent.

$r \in \mathbb{Q}$. (This depends on the density of E.) Then define $E_0 \subset Y$ to consist of exactly one element from each of the sets $B_r(x) \cap Y$ for $x \in E$ and $r \in \mathbb{Q}$, whenever those sets are nonempty. Show that E_0 is countable and dense in Y.

2.3. The definition of a manifold. Hopefully you now have sufficient motivation to accept the following definition.

DEFINITION 2.18. Assume $k \in \mathbb{N} \cup \{\infty\}$. A **differentiable manifold of class C^k** (*differenzierbare Mannigfaltigkeit von der Klasse C^k*) or **C^k-manifold** (*C^k-Mannigfaltigkeit*) is a set M endowed with a C^k-structure (see Definition 2.6) such that the induced topology on M is metrizable and separable. In the case $k = \infty$, we also call M a **smooth manifold** (*glatte Mannigfaltigkeit*). We say that M is **n-dimensional** and refer to M as an **n-manifold**, written

$$\dim M = n,$$

if every chart in its differentiable structure is n-dimensional.[14]

REMARK 2.19. For the purposes of this course, you are essentially free to ignore the separability condition in Definition 2.18, as nothing in our study of differential geometry will truly depend on it. An example of something that satisfies every condition in the definition except separability would be the disjoint union of *uncountably* many copies of a manifold (see §2.4.3 below for more on disjoint unions); in fact, one can show that the condition on separability in our definition is equivalent to requiring M to have at most countably many connected components. One does sometimes need to know this for important results in differential *topology*, e.g. there is a theorem guaranteeing that every smooth n-manifold M can be embedded as a smooth submanifold of $\mathbb{R}^{2n+1}$, and this would clearly contradict Exercise 2.17 if M were not separable. (This issue is related to the second countability axiom—see Remark 2.21.)

REMARK 2.20. If you prefer never to think about topological spaces, then you can read Definition 2.18 as saying that a manifold M is a separable metric space endowed with an atlas $\{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ for which the sets $\mathcal{U}_\alpha \subset M$ are open and the bijections $x_\alpha : \mathcal{U}_\alpha \to x_\alpha(\mathcal{U}_\alpha) \subset \mathbb{R}^n$ are continuous with continuous inverses. Calling M a "metric space" comes however with the following caveat: while the *existence* of a suitable metric on M is an important condition, the *choice* of metric on M is not considered a part of its intrinsic structure, i.e. you are free to replace it with any other metric that has the above properties with respect to the atlas. This is why we have used the word "metrizable" in Definition 2.18 instead of just calling M a "metric space".

REMARK 2.21. For students who have seen some topology, the more standard definition of a manifold found in many textbooks would replace the conditions of metrizability and separability with the conditions that M is *Hausdorff* and *second countable*. This gives an equivalent definition, though proving this equivalence would require more of a digression into point-set topology than we have space for here; the details can (mostly) be found in [**Lee11**, Chapter 2].

REMARK 2.22. Another reasonable guess for a good definition of a manifold would be to drop metrizability and separability from Definition 2.18 but still require M to be Hausdorff (thus excluding things like Example 2.16). It turns out that this also does not include enough conditions to rule out some pathological behavior. The issue here is that a locally Euclidean Hausdorff space may fail to be *paracompact*, in which case the construction of basic geometric objects like Riemannian metrics becomes impossible. (We will discuss paracompactness and its applications later in the course.) If you have some topological background and would like to see some examples

[14]Note that in our general definition of a manifold, M might admit multiple charts of different dimensions. One can show however that each individual connected component of M is itself a manifold with a uniquely defined dimension. For this reason we will usually only consider manifolds that have a well-defined dimension.

of the kinds of pathological behavior I'm talking about, see the discussion of the *long line* and *Prüfer surface* in [**Wen23**, Lecture 18].

In this course, we will almost always consider only the case $k = \infty$ of Definition 2.18, so that we speak of *smooth* manifolds. Actually, a large portion of differential geometry still makes sense for C^1-manifolds, though the important notion of *curvature* on a Riemannian manifold depends on second derivatives of the metric, and thus only makes sense on manifolds of class C^2. In either case, one has to be very careful in every proof so as not to differentiate anything more times than is allowed, and since the most important examples of manifolds are of class C^∞, it is conventional to avoid this annoyance by restricting attention to the smooth case. There is an additional reason to allow this restriction: according to a standard theorem in differential topology (see [**Hir94**, Theorem 2.9]), every manifold of class C^1 can be made into a *smooth* manifold by removing some of the charts in its maximal C^1-atlas. In this sense, one does not lose any significant generality by ignoring manifolds that are differentiable but not smooth.

You may have noticed on the other hand that Definition 2.18 also makes sense for $k = 0$, though in this case one cannot use the word "differentiable"; manifolds of class C^0 are called **topological manifolds** (*topologische Mannigfaltigkeiten*). These really are a different beast than differentiable manifolds: for every $n \geqslant 4$, there exist topological n-manifolds that *do not admit* any differentiable structure, i.e. their topology is not compatible with any atlas of class C^k for $k \geqslant 1$. Proving such things typically requires very advanced techniques, e.g. from mathematical gauge theory, which uses nonlinear PDEs to derive topological restrictions on smooth manifolds. (The classic introduction to this subject is [**DK90**].) In any case, the study of topological manifolds as such belongs squarely to the subject of topology, not differential geometry, so we will say no more about it here.

2.4. Some basic examples.

2.4.1. *Vector spaces.* For each integer $n \geqslant 0$, $\mathbb{R}^n$ admits a canonical smooth atlas consisting of a single n-dimensional chart, namely $(\mathbb{R}^n, \mathrm{Id})$. The smoothness of this atlas is a triviality: since there is only one chart, there is only one transition map to consider, which is the identity map and is therefore smooth. The unique extension of this atlas to a maximal smooth atlas on $\mathbb{R}^n$ defines what we will call the **standard smooth structure** on $\mathbb{R}^n$. The topology induced by this atlas is the standard one, which can also be defined in terms of the standard Euclidean metric; this follows via Remark 2.14 from the observations that $\mathbb{R}^n \subset \mathbb{R}^n$ is an open subset and $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$ is a homeomorphism. It follows that $\mathbb{R}^n$ with its standard smooth structure is metrizable and (in light of the countable dense subset $\mathbb{Q}^n \subset \mathbb{R}^n$) separable. We conclude that $\mathbb{R}^n$ is, in a natural way, a smooth n-dimensional manifold. Note that it is *possible* to define different smooth structures on $\mathbb{R}^n$, as shown by Example 2.7 in the case $n = 1$, but whenever we discuss $\mathbb{R}^n$ as a manifold in this course, we will always assume unless stated otherwise that it carries its standard smooth structure.

Since every real n-dimensional vector space V is isomorphic to $\mathbb{R}^n$, one can always choose such an isomorphism $\Phi : V \to \mathbb{R}^n$ and similarly regard V as a smooth n-manifold with an atlas consisting of the global chart (V, Φ). While the choice of isomorphism Φ here is typically not canonical, the resulting smooth structure on V is, since any other choice of isomorphism $\Psi : V \to \mathbb{R}^n$ would produce a chart (V, Ψ) that is related to (V, Φ) by the transition map $\Phi \circ \Psi^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. The latter is a vector space isomorphism, and thus a smooth map with a smooth inverse. In this way, we can regard every real n-dimensional vector space naturally as a smooth n-manifold.

2.4.2. *Open subsets.* If M is an n-dimensional C^k-manifold with atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$, then any open subset $\mathcal{O} \subset M$ admits a natural atlas

$$\mathcal{A}_{\mathcal{O}} := \{(\mathcal{U}_\alpha \cap \mathcal{O}, x_\alpha|_{\mathcal{U}_\alpha \cap \mathcal{O}})\}_{\alpha \in I},$$

which is also of class C^k since its transition maps are all restrictions of transition maps from $\mathcal{A}$ to open subsets. The key point here is that since $\mathcal{O} \subset M$ is open, each $\mathcal{U}_\alpha \cap \mathcal{O}$ is an open subset of $\mathcal{U}_\alpha$ and is thus mapped homeomorphically by x_α to another open subset of $\mathbb{R}^n$, making it an n-dimensional chart on $\mathcal{O}$. This atlas endows $\mathcal{O}$ with a natural C^k-structure, and since it is a subset of a separable metrizable space, Exercise 2.17 implies that it is also separable and metrizable, and is thus an n-dimensional C^k-manifold. Combining this with §2.4.1, we can now regard every open subset of $\mathbb{R}^n$ as a smooth n-manifold in a natural way.

2.4.3. *Disjoint unions.* The **disjoint union** (*disjunkte Vereinigung*) of a collection of sets $\{M_j\}_{j\in J}$ can be defined as the set

$$\coprod_{j\in J} M_j := \left\{(j,t) \;\middle|\; j \in J,\ t \in M_j\right\}.$$

Here J can be an arbitrary index set: finite, countable or uncountable. In the special case where J is finite, e.g. if $J = \{1,\dots,N\}$, we also use the notation

$$M_1 \amalg \ldots \amalg M_N := \coprod_{j=1}^N M_j := \coprod_{j\in\{1,\dots,N\}} M_j.$$

Identifying each of the individual sets M_j with the subset $\{j\} \times M_j \subset \coprod_{j\in J} M_j$, we can think of $\coprod_{j\in J} M_j$ as literally a union of all the sets M_j, with the caveat that for $j \ne k$, M_j and M_k are always *disjoint* as subsets of $\coprod_{j\in J} M_j$, even if as abstract sets they have elements in common. For example, the set $S^1 \amalg S^1$ contains two copies of every point on the circle, and is thus not the same set as $S^1 \cup S^1 = S^1$. If you think of S^1 as the unit circle in $\mathbb{R}^2$, then the definition above gives $S^1 \amalg S^1 = \{1,2\} \times S^1 \subset \mathbb{R}^3$, so the disjoint union consists of two copies of the circle that live in disjoint planes in $\mathbb{R}^3$.

Suppose now that each of the sets M_j is a C^k-manifold with atlas $\mathcal{A}^{(j)} = \{(\mathcal{U}_\alpha^{(j)}, x_\alpha^{(j)})\}_{\alpha\in I_j}$. Regarding each set M_j as a subset of $\coprod_{j\in J} M_j$ makes each of the sets $\mathcal{U}_\alpha^{(j)}$ also into subsets of $\coprod_{j\in J} M_j$, such that $\mathcal{U}_\alpha^{(j)} \cap \mathcal{U}_\beta^{(k)} = \varnothing$ whenever $j \ne k$. It follows that the union

$$\mathcal{A} := \bigcup_{j\in J} \mathcal{A}^{(j)}$$

defines an atlas of class C^k on $\coprod_{j\in J} M_j$, whose set of transition maps is just the union of the sets of transition maps for all the atlases $\mathcal{A}^{(j)}$. (Transition maps relating two charts $(\mathcal{U}_\alpha^{(j)}, x_\alpha^{(j)})$ with $(\mathcal{U}_\beta^{(k)}, x_\beta^{(k)})$ with $j \ne k$ do not arise here since their overlap is always empty.)

It does *not* follow however that every disjoint union of a collection of C^k-manifolds is naturally a C^k-manifold—this is one of the few situations where we have to pay attention to the condition of separability. The topology induced by the atlas $\mathcal{A}$ on $\coprod_{j\in J} M_j$ is the so-called **disjoint union topology**, in which a subset $\mathcal{O} \subset \coprod_{j\in J} M_j$ is open if and only if $\mathcal{O} \cap M_j$ is an open subset of M_j for every $j \in J$. If the sets M_j are nonempty for uncountably many distinct values of $j \in J$, then no *countable* subset $E \subset \coprod_{j\in J} M_j$ can have an element in every one of the subsets M_j, and it follows that E cannot be dense, so the disjoint union cannot be separable. On the other hand, one can show (see Exercise 2.23 below) that every *finite* or *countable* disjoint union of separable metrizable spaces is also separable and metrizable. We conclude that for any $N \in \mathbb{N} \cup \{\infty\}$ and any finite or countable collection $\{M_j\}_{j=1}^N$ of C^k-manifolds, the disjoint union $\coprod_{j=1}^N M_j$ is also a C^k-manifold in a natural way. Moreover, if $\dim M_j = n$ for every j, then the disjoint union is also n-dimensional.

EXERCISE 2.23.

(a) Show that for any metric space (X, d), the formula

$$d'(x, y) := \begin{cases} d(x, y) & \text{if } d(x, y) < 1, \\ 1 & \text{if } d(x, y) \geqslant 1 \end{cases}$$

defines another metric d' on X that induces the same topology as d.

(b) Show that for any collection of metric spaces $\{(X_j, d_j)\}_{j \in J}$ with $d_j(x, y) \leqslant 1$ for all $j \in J$ and $x, y \in X_j$, the formula

$$d(x, y) := \begin{cases} d_j(x, y) & \text{if } x, y \in X_j \text{ for some } j \in J, \\ 2 & \text{if } x \in X_j \text{ and } y \in X_k \text{ for some } j, k \in J \text{ with } j \neq k \end{cases}$$

defines a metric on $\coprod_{j \in J} X_j$ that induces the disjoint union topology.

(c) Show that the metric d on $\coprod_{j \in J} X_j$ in part (b) is separable if J is a finite or countable set and all of the metric spaces (X_j, d_j) are separable.

EXERCISE 2.24. Recall that a metrizable space[15] is called **compact** (*kompakt*) if every open covering has a finite subcover. Show that a disjoint union $\coprod_{j \in J} M_j$ is compact if and only if J is finite and M_j is compact for every $j \in J$.

2.4.4. *Dimension zero.* You may not have thought about the case $n = 0$ when we defined the notion of an n-dimensional chart, but the definition in that case does make sense: $\mathbb{R}^0$ consists of a single point, and its only nontrivial open subset is itself, so if $(\mathcal{U}, x)$ is a 0-dimensional chart on M, then $\mathcal{U} \subset M$ is a single point. It follows that if M is a 0-dimensional manifold with atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$, then every point of M is its own open set, implying that *every* subset of M is open. This is known as the **discrete topology**, and it is always metrizable; a suitable metric is the **discrete metric**, defined by

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

The only dense subset of M in this topology is M itself, so separability requires M to be finite or countable. We conclude: a 0-dimensional manifold is simply a *finite or countable discrete set*, and it is compact if and only if it is finite. Equivalently, every 0-dimensional manifold can be identified with the disjoint union of at most countably many copies of the manifold $\mathbb{R}^0$, which is a single point. Notice that since every map from $\mathbb{R}^0$ to itself is trivially smooth, every atlas on a 0-manifold is automatically a smooth atlas.

2.4.5. *Dimension one.* We have seen two explicit examples thus far of 1-dimensional manifolds: $\mathbb{R}$ and S^1, where the former carries its standard smooth structure as defined in §2.4.1, and the latter has a smooth structure that we defined using two charts based on polar coordinates in Lecture 1. We can now add to this list arbitrary open subsets of each, and arbitrary finite or countable disjoint unions of such open subsets. In this entire list, the only actual *compact* examples are S^1 and its finite disjoint unions; the compactness of the circle $S^1 \subset \mathbb{R}^2$ follows from the general fact that closed and bounded subsets of Euclidean space are compact. Up to a natural notion of equivalence for smooth manifolds that we will discuss in the next lecture, it turns out that these really are the *only* examples: in particular, every compact and *connected* 1-manifold is "diffeomorphic" to S^1. Later when we discuss manifolds with boundary, we will have to add the compact interval $[0, 1]$ to the list of compact 1-manifolds up to diffeomorphism. Similarly, it turns out that every noncompact

[15] In fact this definition is also valid for arbitrary topological spaces.

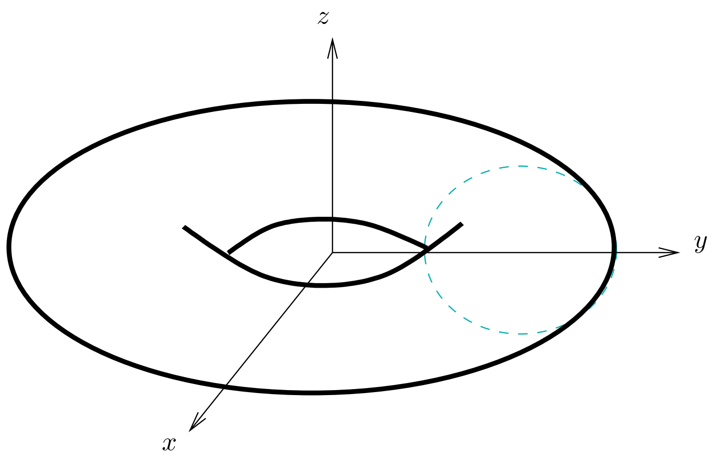


FIGURE 5. A representation of the torus $\mathbb{T}^2$ as a submanifold of $\mathbb{R}^3$.

connected 1-manifold is diffeomorphic to $\mathbb{R}$. We will not prove such classification results in this course, nor make use of them, but the curious reader will find a sketch of the corresponding result about connected topological 1-manifolds up to homeomorphism in [**Wen23**, Lecture 18]. Note that this is one of the important results that becomes false if one drops the metrizability condition from the definition of a manifold; we already saw one peculiar counterexample in Example 2.16, and another is the so-called "long line", which is essentially a union of *uncountably* many compact intervals glued together at their end points (see [**Wen23**, Lecture 18] or [**Spi99**, Appendix to Chapter 1]).

2.4.6. *Cartesian products.* Since we have no plans to discuss infinite-dimensional manifolds in this course, we will not talk about infinite products, but finite products still provide a useful way of producing new manifolds from old ones. Assume M and N are C^k-manifolds of dimensions m and n respectively, with atlases $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ on M and $\mathcal{B} = \{(\mathcal{V}_\beta, y_\beta)\}_{\beta \in J}$ on N. For each $(\alpha, \beta) \in I \times J$, one can then define a **product chart** on $M \times N$ with domain $\mathcal{U}_\alpha \times \mathcal{V}_\beta$ by

$$\mathcal{U}_\alpha \times \mathcal{V}_\beta \to \mathbb{R}^{m \times n} : (p, q) \mapsto (x_\alpha(p), y_\beta(q)).$$

Each of the transition maps relating two product charts is just the Cartesian product of a transition map from $\mathcal{A}$ with one from $\mathcal{B}$, thus they are all of class C^k, and the collection of all product charts therefore defines an atlas of class C^k and makes $M \times N$ into a C^k-manifold of dimension $m + n$.[16] One can of course repeat this construction finitely many times to make any finite product of manifolds $M_1 \times \ldots \times M_N$ into a manifold.

An important special case of this construction is the compact smooth n-manifold known as the n-**torus**, defined by

$$\mathbb{T}^n := \underbrace{S^1 \times \ldots \times S^1}_{n}.$$

In the case $n = 1$, this is just another name for the circle, but the most popular torus is the case $n = 2$: as we've defined it, $\mathbb{T}^2$ is literally a subset of $\mathbb{R}^4$, but for visualization purposes there is also a favorite way of embedding it in $\mathbb{R}^3$, as shown in Figure 5.

[16]Note that even if $\mathcal{A}$ and $\mathcal{B}$ are maximal atlases, the set of all product charts is generally not maximal, but this is immaterial since it has a unique maximal extension.

The n-torus for $n \geqslant 3$ is less straightforward to visualize, but it is often useful to think of it[17] as the quotient of $\mathbb{R}^n$ by the lattice $\mathbb{Z}^n$, using the bijection

$$\mathbb{R}^n/\mathbb{Z}^n \to \underbrace{S^1 \times \ldots \times S^1}_{n} : [(\theta_1, \ldots, \theta_n)] \mapsto (e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_n}),$$

where for computational convenience we have replaced $\mathbb{R}^2$ with $\mathbb{C}$ in order to describe points in the unit circle S^1 as complex exponentials. Under this identification, a point in $\mathbb{T}^n$ is represented by a vector in $\mathbb{R}^n$, with the understanding that two vectors represent the same point in the torus if and only if they differ by a vector with integer coordinates. This perspective is especially useful in the study of Fourier series, as a function $f : \mathbb{R}^n \to \mathbb{C}$ that is 1-periodic in each of the n variables can now be regarded equivalently as a function $f : \mathbb{T}^n \to \mathbb{C}$.

Exercise 2.25. Convince yourself that the natural smooth structure on $\underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n}$ derived from the standard smooth structure of $\mathbb{R}$ is the same as the standard smooth structure of $\mathbb{R}^n$.

2.4.7. *The projective plane and the Klein bottle.* We conclude with two explicit examples of surfaces (i.e. smooth 2-manifolds) that are somewhat harder to visualize, because they cannot be embedded in $\mathbb{R}^3$.[18]

The **projective plane** (*projektive Ebene*) is the set of equivalence classes

$$\mathbb{RP}^2 := S^2/\sim,$$

where the equivalence relation is defined by $p \sim p$ and $p \sim -p$ for all $p \in S^2 \subset \mathbb{R}^3$, meaning that every point p in the unit sphere gets identified with its *antipodal* point $-p$. (For more on why this might be a natural object to define, see Exercise 2.26 below.) If you have ever been on a long-haul international flight, then you are familiar with the notion of traversing a continuous path along S^2. In order to picture a continuous path on $\mathbb{RP}^2$, you should imagine that there are always two identical and interchangeable airplanes, containing identical copies of the same crews and passengers, constrained to fly at exact antipodal points over the Earth. If one of those airplanes flies from Shanghai to Buenos Aires while the other one flies along the antipodal path,[19] then since the two planes are completely interchangeable, they can be understood to describe a *closed loop* on $\mathbb{RP}^2$. Got it? Good.

It is relatively easy to see that $\mathbb{RP}^2$ is a smooth 2-manifold in a natural way. First, it has a natural metric, in which one can describe each point of $\mathbb{RP}^2$ as a set consisting of two points in S^2 and define the distance between two points in $\mathbb{RP}^2$ as the distance between those two sets. The fact that S^2 is separable (as a subset of the separable metric space $\mathbb{R}^3$) implies easily that $\mathbb{RP}^2$ is also separable. One can also derive a smooth atlas on $\mathbb{RP}^2$ from the one that we already constructed on S^2 in Exercise 1.7: the only issue is that some of the charts need to have their domains shrunk so that they no longer contain any pairs of antipodal points, as the coordinate map will otherwise fail to be injective, but this can easily be done.

The second example is the **Klein bottle** (*Kleinsche Flasche*), a picture of which is shown in Figure 6. The picture must be interpreted with caution, since what it shows is not really a manifold in the usual sense, but if you imagine perturbing part of it in an unseen fourth dimension so that

[17] In fact, many sources in the literature prefer to *define* $\mathbb{T}^n$ as the quotient group $\mathbb{R}^n/\mathbb{Z}^n$, in which case its smooth structure can be derived from the standard smooth structure of $\mathbb{R}^n$ using a general result about quotients by discrete group actions.

[18] The claim that embedding them into $\mathbb{R}^3$ is *impossible* is something I expect you to find plausible, but not obvious. Proving it would require some methods from topology which are beyond the scope of this course.

[19] According to the British science fiction TV series *Torchwood*, Buenos Aires and Shanghai are at exact antipodal points on the Earth. Wikipedia says this is true up to an error of about 400km. Let's just pretend it's true.

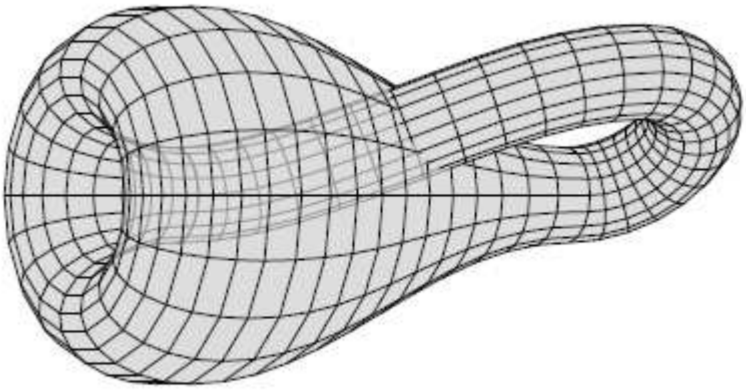


FIGURE 6. An immersion of the Klein bottle into $\mathbb{R}^3$. It is not an embedding because it intersects itself. (We will discuss the precise meanings of the words "immersion" and "embedding" in Lecture 4.)

part of the surface no longer has to pass through another part, then you get the right intuition. The picture also shows a "grid" structure similar to the coordinate grid one would obtain on $\mathbb{T}^2$ after identifying it with $\mathbb{R}^2/\mathbb{Z}^2$, but the Klein bottle is not the same thing as the torus. The latter can be identified with the quotient

$$(\mathbb{R} \times (\mathbb{R}/\mathbb{Z})) \Big/ \sim$$

by the smallest equivalence relation on $\mathbb{R} \times (\mathbb{R}/\mathbb{Z})$ such that $(s, [t]) \sim (s+1, [t])$ for all $s, t \in \mathbb{R}$. One obtains a rigorous definition of the Klein bottle from this via a reversal of orientation: instead of $(s, [t]) \sim (s+1, [t])$, one takes the smallest equivalence relation on $\mathbb{R} \times (\mathbb{R}/\mathbb{Z})$ such that

$$(s, [t]) \sim (s+1, [-t])$$

for all $s, t \in \mathbb{R}$. If you think about what grid lines of the form $\{s = \text{const}\}$ and $\{t = \text{const}\}$ look like in the set of equivalence classes defined via this relation, you will end up with something resembling Figure 6. It is not difficult to construct an atlas of smoothly compatible 2-dimensional charts on this quotient: the basic idea is to view it as a quotient of $\mathbb{R}^2$, and restrict the canonical global chart of $\mathbb{R}^2$ to neighborhoods that are sufficiently small so as to contain at most one element from every equivalence class.

EXERCISE 2.26. The projective plane is the $n = 2$ case of the **real projective n-space** (*reeller projektiver Raum*)

$$\mathbb{RP}^n := S^n \big/ \sim,$$

where here again the equivalence relation identifies antipodal points $x \sim -x \in S^n \subset \mathbb{R}^{n+1}$. A useful interpretation of this definition comes from the observation that there is a unique line through the origin passing through each pair of points $\{x, -x\} \subset \mathbb{R}^{n+1}$. One can therefore view $\mathbb{RP}^n$ equivalently as the space of all *lines through the origin in* $\mathbb{R}^{n+1}$, which can be defined more precisely as the quotient

$$\mathbb{RP}^n = (\mathbb{R}^{n+1} \backslash \{0\}) \big/ \sim$$

where two nontrivial vectors $v, w \in \mathbb{R}^{n+1}$ are now considered equivalent if and only if $v = \lambda w$ for some $\lambda \in \mathbb{R}$. From this perspective, it is convenient to denote points in $\mathbb{RP}^n$ via so-called

homogeneous coordinates, in which the symbol

$$[x_0 : \ldots : x_n] \in \mathbb{RP}^n$$

means the equivalence class containing the vector $(x_0, \ldots, x_n) \in \mathbb{R}^{n+1}\backslash\{0\}$.

The homogeneous coordinates can be used to define an explicit smooth atlas on $\mathbb{RP}^n$. For $j = 0, \ldots, n$, define

$$\mathcal{U}_j := \left\{ [x_0 : \ldots : x_n] \in \mathbb{RP}^n \;\middle|\; x_j \neq 0 \right\}$$

and a map $\varphi_j : \mathbb{R}^n \to \mathbb{RP}^n$ by

$$\varphi_j(t_1, \ldots, t_n) := [t_1 : \ldots : t_j : 1 : t_{j+1} : \ldots : t_n].$$

Show that φ_j is an injective map onto $\mathcal{U}_j$, so $(\mathcal{U}_j, \varphi_j^{-1})$ is a chart, and compute the transition maps relating any two of the charts constructed in this way for different values of $j = 0, \ldots, n$. Show that these $n + 1$ charts together form a smooth atlas.

3. Smooth maps and tangent vectors

We have several more definitions to get through before the subject of differential geometry gets seriously underway. In this lecture we clarify what it means for a map between two manifolds to be differentiable, and what kind of object its derivative is.

3.1. Smooth maps between manifolds. We defined in §2.1 what it means for a real-valued function on a smooth manifold to be smooth (see Definition 2.3). The following is based on the same idea.

Definition 3.1. Assume M and N are manifolds of dimensions m and n respectively, with differentiable structures $\mathcal{A}_M$ and $\mathcal{A}_N$ of class C^k. A continuous map $f : M \to N$ is said to be **of class** C^r for some $r \leqslant k$ (or **smooth** in the case $r = k = \infty$) if for every pair of charts $(\mathcal{U}, x) \in \mathcal{A}_M$ and $(\mathcal{V}, y) \in \mathcal{A}_N$, the map

$$\mathbb{R}^m \overset{\text{open}}{\supset} x(\mathcal{U} \cap f^{-1}(\mathcal{V})) \xrightarrow{y \circ f \circ x^{-1}} y(\mathcal{V}) \overset{\text{open}}{\subset} \mathbb{R}^n$$

is of class C^r.

In other words, a map $f : M \to N$ is of class C^r if it looks like a map of class C^r when expressed in local coordinates on both the domain and the target. The assumption $r \leqslant k$ is again crucial here, and guarantees that for any given point $p \in M$, the question of whether f is of class C^r near p does not depend on the charts one has to choose near $p \in M$ and $f(p) \in N$. Note that we had to explicitly assume f was continuous in this definition: this assumption guarantees that $f^{-1}(\mathcal{V}) \subset M$ is an open set, so that $x(\mathcal{U} \cap f^{-1}(\mathcal{V}))$ is open in $\mathbb{R}^n$, and differentiability on this domain can therefore be checked.

The set of C^k maps from M to N is often denoted by

$$C^k(M, N) = \left\{ f : M \to N \;\middle|\; f \text{ is of class } C^k \right\}.$$

One can endow this space with various natural topologies to make it into a topological (and sometimes also metrizable) space, though you should be aware that it is generally not a vector space, since N is not. On the other hand, the special case $N = \mathbb{R}$ is quite important, and is often abbreviated

$$C^k(M) := C^k(M, \mathbb{R}).$$

This *is* a vector space in a natural way, i.e. real-valued functions on a manifold M can be added and multiplied by constants.

EXERCISE 3.2. Show that for the standard smooth structure on $\mathbb{R}$ defined in §2.4.1, the notion of differentiability for a map $f : M \to \mathbb{R}$ as given in Definition 3.1 matches our previous definition for real-valued functions (Definition 2.3).

Up until this point I have been including non-smooth manifolds in the picture. I could continue doing this, but it would require frequently including slightly annoying extra hypotheses (like $r \leqslant k$) in statements of results, and the generality one gains by doing this does not fully compensate for the annoyance, so I will mostly assume $k = \infty$ from now on.

We can now define the natural notion of equivalence for smooth manifolds.

DEFINITION 3.3. For two smooth manifolds M and N, a smooth map $f : M \to N$ is called a **diffeomorphism** (*Diffeomorphismus*) if it is bijective and its inverse $f^{-1} : N \to M$ is also smooth. Two smooth manifolds are called **diffeomorphic** (*diffeomorph*) if there exists a diffeomorphism between them.

EXERCISE 3.4. Viewing S^1 as the unit circle in $\mathbb{C}$, the quotient group $\mathbb{R}^n/\mathbb{Z}^n$ admits a natural bijection to the n-torus $\mathbb{T}^n = S^1 \times \ldots \times S^1$, given by

$$\mathbb{R}^n/\mathbb{Z}^n \to \mathbb{T}^n : [(\theta_1, \ldots, \theta_n)] \mapsto (e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_n}).$$

For each $v \in \mathbb{R}^n$, choose a neighborhood $\widetilde{\mathcal{U}}_v \subset \mathbb{R}^n$ of v that is small enough to contain at most one element from each equivalence class in $\mathbb{R}^n/\mathbb{Z}^n$, and use this to define an n-dimensional chart $(\mathcal{U}_v, x_v)$ of the form

$$\mathcal{U}_v = \left\{ [w] \in \mathbb{R}^n/\mathbb{Z}^n \;\middle|\; w \in \widetilde{\mathcal{U}}_v \right\}, \qquad x_v([w]) = w.$$

Show that the collection of all charts of this form determines a smooth atlas on $\mathbb{R}^n/\mathbb{Z}^n$ such that the bijection to $\mathbb{T}^n$ described above is a diffeomorphism.

3.2. Tangent and cotangent spaces. Let us start this discussion with a concrete example: on the unit sphere $S^2 \subset \mathbb{R}^3$, a *tangent vector* to S^2 at a point $p \in S^2$ is by definition any vector of the form

$$\gamma'(0) \in \mathbb{R}^3,$$

where $\gamma : (-\epsilon, \epsilon) \to S^2$ is any choice of smooth path in $\mathbb{R}^3$ whose image is in S^2 and satisfies $\gamma(0) = p$. It should be easy to convince yourself that the set of all vectors of this form is a linear subspace of $\mathbb{R}^3$, namely, it is the orthogonal complement of p. We would now like to generalize this notion to an arbitrary smooth manifold, without needing to assume that is a subset of some Euclidean space.

For the rest of this subsection, assume M is a smooth manifold and $p \in M$. Having defined what a smooth map between manifolds is, we can fix the standard smooth structure on small intervals such as $(-\epsilon, \epsilon) \subset \mathbb{R}$ and talk about smooth maps $\gamma : (-\epsilon, \epsilon) \to M$. If $\gamma(0) = p \in M$, then we will refer to any such smooth map as a **path through p in M**. Note that the value of $\epsilon > 0$ here is not fixed, so it is allowed to be arbitrarily small.

Let us say that two paths α, β through p in M are **tangent** if for some some chart $(\mathcal{U}, x)$ with $p \in \mathcal{U}$,

$$\frac{d}{dt}(x \circ \alpha)\bigg|_{t=0} = \frac{d}{dt}(x \circ \beta)\bigg|_{t=0}.$$

It is easy to show that this condition does not depend on the choice of chart: indeed, if $(\mathcal{V}, y)$ is another chart with $p \in \mathcal{V}$, then for all t close enough to 0 so that $\alpha(t) \in \mathcal{U} \cap \mathcal{V}$, we have $(y \circ \alpha)(t) = (y \circ x^{-1}) \circ (x \circ \alpha)(t)$ and thus by the chain rule,

$$(y \circ \alpha)'(0) = D(y \circ x^{-1})(x(p))(x \circ \alpha)'(0), \tag{3.1}$$

where $D(y \circ x^{-1})(x(p)) : \mathbb{R}^n \to \mathbb{R}^n$ denotes the derivative of the transition map $y \circ x^{-1}$ at $x(p)$, which is an invertible linear map since $y \circ x^{-1}$ is smooth and has a smooth inverse. Since $(y \circ \beta)'(0)$ is related to $(x \circ \beta)'(0)$ in the same way, it is equal to $(y \circ \alpha)'(0)$ if and only if $(x \circ \beta)'(0) = (x \circ \alpha)'(0)$.

Definition 3.5. A **tangent vector** (*Tangentialvektor*) to M at p is an equivalence class $[\gamma]$ of paths γ through p in M, where two paths are considered equivalent if and only if they are tangent. The set of all tangent vectors to M at p is called the **tangent space** (*Tangentialraum*) to M at p, and is denoted by

$$T_pM = \{[\gamma] \mid \gamma \text{ a path through } p \text{ in } M\}.$$

This definition of T_pM has many intuitive advantages, but it leaves several details unclear, foremost among them the fact that T_pM is a vector space. In order to see this, we'll need to make more use of coordinates.

Proposition 3.6. *The tangent space T_pM has a unique vector space structure such that for any smooth n-dimensional chart $(\mathcal{U}, x)$ with $p \in \mathcal{U}$, the map*

$$d_px : T_pM \to \mathbb{R}^n : [\gamma] \mapsto (x \circ \gamma)'(0) \tag{3.2}$$

is a vector space isomorphism. In particular, every tangent space of a smooth n-manifold is naturally an n-dimensional vector space.

Proof. The map (3.2) is a bijection by definition, so one can clearly always *choose* a chart $(\mathcal{U}, x)$ and define a vector space structure on T_pM so as to make this map an isomorphism. The point is then to show that any other choice of chart $(\mathcal{V}, y)$ would have given the same vector space structure on T_pM. This follows from the formula

$$d_py \circ (d_px)^{-1} = D(y \circ x^{-1})(x(p)) : \mathbb{R}^n \to \mathbb{R}^n,$$

which follows from (3.1) and shows that this transformation is itself a vector space isomorphism. □

Example 3.7. If M is an open subset of an n-dimensional vector space V, then the derivative $\gamma'(0)$ for a smooth path $\gamma : (-\epsilon, \epsilon) \to V$ can be defined in the classical way as a vector in V, giving rise to a canonical map

$$T_pM \to V : [\gamma] \mapsto \gamma'(0)$$

for every $p \in M$. It is a straightforward exercise to show that this map is a vector space isomorphism.

In the future, we shall always use this isomorphism to identify tangent spaces on open subsets of a vector space V with V itself, so that we do not need to talk about equivalence classes of paths. In particular, every tangent space on an open subset of $\mathbb{R}^n$ is in this way canonically identified with $\mathbb{R}^n$. We will see in §4.3 below that whenever N is a submanifold of M, one can also naturally regard T_pN for each $p \in N$ as a linear subspace of T_pM, so in the special case where N is a submanifold of $\mathbb{R}^n$, its tangent spaces will all naturally be subspaces of $\mathbb{R}^n$. This means that for the vast majority of examples we are interested in, it will not be necessary to use the original definition in terms of equivalence classes of paths for describing a tangent space.

Exercise 3.8. Show that for two smooth manifolds M, N and any two points $p \in M$ and $q \in N$, there is a canonical vector space isomorphism $T_{(p,q)}(M \times N) = T_pM \times T_qN$.

In linear algebra, it is often useful to associate to any vector space V its **dual space** (*Dualraum*), which is the space of all scalar-valued linear maps on V. Assuming V is a *real* (rather than complex) vector space, this can be denoted by

$$V^* := \text{Hom}(V, \mathbb{R}),$$

where for two real vector spaces V, W in general we denote by $\mathrm{Hom}(V, W)$ the vector space of linear maps $V \to W$. When V is a tangent space T_pM on a manifold M, its dual space is called the **cotangent space** (*Kotangentialraum*) to M at p and denoted by

$$T_p^*M := \mathrm{Hom}(T_pM, \mathbb{R}).$$

Its elements are called **cotangent vectors** (*Kotangentialvektoren*), or sometimes also **covectors**.

REMARK 3.9. Among physicists, covectors are often called "covariant vectors", while ordinary tangent vectors are called "contravariant vectors". I will not use this terminology.

3.3. The tangent bundle. The usefulness of the following definition will probably not be obvious to you at first glance, but it will become more apparent when we start differentiating smooth maps.

DEFINITION 3.10. The **tangent bundle** (*Tangentialbündel*) TM of a smooth manifold M is the union of all its tangent spaces:

$$TM := \bigcup_{p \in M} T_pM.$$

The map $\pi : TM \to M$ such that $\pi^{-1}(p) = T_pM \subset TM$ for each $p \in M$ is called the **tangent projection**, and the subset in TM consisting of the zero vectors $0 \in T_pM$ for all $p \in M$ is called the **zero-section** (*Nullschnitt*) of TM. As subsets of TM, the individual tangent spaces $T_pM \subset TM$ for each $p \in M$ are sometimes referred to as the **fibers** (*Fasern*) of the tangent bundle.

Note that for distinct points $p \neq q \in M$, the tangent spaces T_pM and T_qM are by definition disjoint sets. Do not be tempted to think that the zero vector in T_pM is the same point as the zero vector in T_qM for $p \neq q$; in fact, there is a natural identification of the zero-section with M, giving rise to a natural inclusion

$$i : M \hookrightarrow TM : p \mapsto 0 \in T_pM. \tag{3.3}$$

At the level of set theory, we could just as well have used the disjoint union notation $\coprod_{p \in M} T_pM$ in Definition 3.10, but we did not do this because it would give a misleading impression about the topology and smooth structure we intend to define on TM.

LEMMA 3.11. *On a manifold M, any n-dimensional chart $(\mathcal{U}, x)$ determines a $2n$-dimensional chart $(T\mathcal{U}, Tx)$ on the tangent bundle TM, where $T\mathcal{U} = \bigcup_{p \in \mathcal{U}} T_pM$ is the tangent bundle of the open subset $\mathcal{U} \subset M$, and $Tx : T\mathcal{U} \to \mathbb{R}^{2n}$ is defined in terms of the linear isomorphism $d_px : T_pM \to \mathbb{R}^n$ of (3.2) by*

$$T\mathcal{U} \supset T_pM \ni X \mapsto (x(p), d_px(X)) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}.$$

If $(\mathcal{V}, y)$ is another chart on M, then transition maps relating the charts $(T\mathcal{V}, Ty)$ and $(T\mathcal{U}, Tx)$ on TM are given by

$$Ty \circ (Tx)^{-1}(q, v) = \left(y \circ x^{-1}(q), D(y \circ x^{-1})(q)v\right).$$

PROOF. The map $Tx : T\mathcal{U} \to \mathbb{R}^{2n}$ is clearly injective, and its image is $x(\mathcal{U}) \times \mathbb{R}^n$, which is open. The stated formula for the transition map $Ty \circ (Tx)^{-1}$ follows from (3.1). □

COROLLARY 3.12. *For any smooth manifold M, the tangent bundle TM can be endowed naturally with the structure of a smooth manifold such that the tangent projection $\pi : TM \to M$, the inclusion $i : M \hookrightarrow TM$ of the zero-section (3.3) and the natural inclusions $T_pM \hookrightarrow TM$ for all $p \in M$ are smooth maps.*[20] *If $\dim M = n$, then $\dim TM = 2n$.*

[20] Here we are using the vector space structure of T_pM to regard it as a smooth manifold as in §2.4.1.

Proof. We endow TM with the unique maximal smooth atlas containing all charts of the form $(T\mathcal{U}, Tx)$ determined via Lemma 3.11 from smooth charts $(\mathcal{U}, x)$ on M.

To check that $\pi : TM \to M$ is a smooth map, one can now write its coordinate expression with respect to any chart $(\mathcal{U}, x)$ on M and the corresponding chart $(T\mathcal{U}, Tx)$ on TM: the resulting map from an open subset of $\mathbb{R}^{2n}$ to $\mathbb{R}^n$ takes the form $(q, v) \mapsto q$, and is thus clearly smooth. Writing down the inclusion of the zero-section $M \hookrightarrow TM$ in similar coordinates produces $q \mapsto (q, 0)$, and for the inclusion $T_pM \hookrightarrow TM$, one obtains $v \mapsto (q, v)$ with $q \in \mathbb{R}^n$ a constant. All of these maps are smooth.

I hope you find it plausible that TM with the atlas constructed above is metrizable and separable. Separability is easy to prove, e.g. one can take the union of countable dense subsets of individual fibers T_pM for all p in some countable dense subset of M, thus forming a countable dense subset of TM. The easiest way I can think of to prove metrizability is by constructing a Riemannian metric on TM, which we will do in Lecture 15. That construction will rely on the assumption that M is metrizable; we will not need to assume this about TM. □

Exercise 3.13. Find a diffeomorphism from the tangent bundle TS^1 to the product manifold $S^1 \times \mathbb{R}$.

One can similarly define a **cotangent bundle** (*Kotangentialbündel*)

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

which satisfies a result analogous to Corollary 3.12. We will postpone the proof of this fact, since it follows from more general results about vector bundles to be discussed later in the course, and we will not really have use for it until then.

3.4. Tangent maps. We can now answer a question you may have wondered about: we know how to define whether a map $f : M \to N$ between manifolds is differentiable, but how does one actually *differentiate* it, i.e. what is its derivative at a point? In the special case $M \overset{\text{open}}{\subset} \mathbb{R}^m$ and $N = \mathbb{R}^n$, the answer you learned from first-year analysis is to view the derivative $Df(p)$ at a point $p \in M$ as a linear map $\mathbb{R}^m \to \mathbb{R}^n$, and according to the chain rule, it satisfies the relation

$$(f \circ \gamma)'(0) = Df(p)\gamma'(0)$$

for any smooth path γ through p. In fact, since any vector in $\mathbb{R}^m$ can be the derivative of some smooth path through p, this formula uniquely characterizes the linear map $Df(p) : \mathbb{R}^m \to \mathbb{R}^n$. It also admits an obvious generalization to the setting of smooth manifolds, using the fact that if $\gamma : (-\epsilon, \epsilon) \to M$ is a path through $p \in M$, then $f \circ \gamma : (-\epsilon, \epsilon) \to N$ is a path through $f(p) \in N$.

Definition 3.14. For two smooth manifolds M, N and a smooth map $f : M \to N$, the **tangent map** (*Tangentialabbildung*) of f is the map

$$Tf : TM \to TN : [\gamma] \mapsto [f \circ \gamma].$$

Its restriction to the tangent space at a specific point $p \in M$ can be denoted by

$$T_pf : T_pM \to T_{f(p)}N,$$

and is also called the **derivative** of f at p.[21]

[21]You will find a variety of alternative notation in the literature for what I am calling T_pf, e.g. $df(p)$ and $Df(p)$ are also popular choices. In these notes, I will try to consistently reserve $Df(p)$ for the notion of derivatives defined in first-year analysis, where one only considers maps between open subsets of Euclidean spaces. The notation df will be reserved for the *differential* of a function valued in $\mathbb{R}$ or another vector space, to be defined in the next lecture.

LEMMA 3.15. *The map $T_p f : T_p M \to T_{f(p)} N$ defined above for a smooth map $f : M \to N$ and a point $p \in M$ is independent of choices, and it is linear. Moreover, if $f : M \to N$ is smooth, then $Tf : TM \to TN$ is also smooth.*

PROOF. All of these statements will become obvious if we write down a local coordinate expression for the map $Tf : TM \to TN$. Choose charts $(\mathcal{U}, x)$ on M and $(\mathcal{V}, y)$ on N with $p \in \mathcal{U}$ and $f(p) \in \mathcal{V}$. These give rise to charts $(T\mathcal{U}, Tx)$ on TM and $(T\mathcal{V}, Ty)$ on TN as in Lemma 3.11, so that given any $[\gamma] \in T_p M$, $Tx([\gamma]) = (x(p), (x \circ \gamma)'(0)) \in \mathbb{R}^m \times \mathbb{R}^m$, and according to the definition of Tf,

$$Ty(Tf([\gamma])) = (y(f(p)), (y \circ (f \circ \gamma))'(0)) \in \mathbb{R}^n \times \mathbb{R}^n.$$

The assumption that f is smooth means that $y \circ f \circ x^{-1}$ is smooth on its domain of definition, which is a neighborhood of $x(p)$ in $\mathbb{R}^m$. On this neighborhood, we can then write $y \circ (f \circ \gamma) = (y \circ f \circ x^{-1}) \circ (x \circ \gamma)$ and apply the chain rule to derive from the above expression,

$$Ty \circ Tf \circ (Tx)^{-1}(x(p), (x \circ \gamma)'(0)) = (y \circ f \circ x^{-1}(x(p)), D(y \circ f \circ x^{-1})(x(p))(x \circ \gamma)'(0)),$$

or if we simplify by writing $q := x(p) \in \mathbb{R}^m$ and $v := (x \circ \gamma)'(0) \in \mathbb{R}^m$,

$$Ty \circ Tf \circ (Tx)^{-1}(q, v) = (y \circ f \circ x^{-1}(q), D(y \circ f \circ x^{-1})(q)v).$$

This formula does not depend on any choice of path γ to represent the tangent vector $[\gamma] \in T_p M$, thus it proves that $Tf([\gamma])$ also does not depend on this choice, and moreover, it defines a smooth map $TM \to TN$ with a linear restriction $T_p M \to T_{f(p)} N$. □

The tangent bundle provides a more elegant language for talking about derivatives than was available in your first-year analysis course. As justification for this claim, I offer the following reformulation of the chain rule in the language of manifolds; it follows directly from the definitions of tangent spaces and tangent maps (which are in themselves crucially dependent on the chain rule from first-year analysis).

PROPOSITION 3.16 (chain rule). *For any pair of smooth maps $f : M \to N$ and $g : N \to Q$ between smooth manifolds, $T(g \circ f) = Tg \circ Tf : TM \to TQ$.* □

COROLLARY 3.17. *If $f : M \to N$ is a diffeomorphism, then so is $Tf : TM \to TN$, and $(Tf)^{-1} = T(f^{-1}) : TN \to TM$.*

PROOF. Observe first that the tangent map to the identity map on M is the identity map on TM. The chain rule then implies $\mathrm{Id}_{TM} = T(f \circ f^{-1}) = Tf \circ T(f^{-1})$. □

REMARK 3.18. Since $T_q\mathbb{R}^n$ is canonically isomorphic to $\mathbb{R}^n$ for every $q \in \mathbb{R}^n$, the tangent bundle $T\mathbb{R}^n$ has a canonical identification with $\mathbb{R}^n \times \mathbb{R}^n$ in which $T_q\mathbb{R}^n = \{q\} \times \mathbb{R}^n$. Under this identification, the chart $Tx : T\mathcal{U} \to \mathbb{R}^n \times \mathbb{R}^n$ on TM derived in Lemma 3.11 from a chart $x : \mathcal{U} \to \mathbb{R}^n$ on M is simply the tangent map of x.

REMARK 3.19. If you are familiar with the language of categories and functors, then you might appreciate the following interpretation of Proposition 3.16. One can define a category Diff whose objects are the smooth manifolds, with morphisms $M \to N$ defined to be smooth maps, hence the isomorphisms in this category are the diffeomorphisms. The construction of the tangent bundle now gives rise to a functor $T : \mathsf{Diff} \to \mathsf{Diff}$ which sends each manifold M to TM and associates to any morphism $f : M \to N$ its tangent map $Tf : TM \to TN$. The formula $T(g \circ f) = Tg \circ Tf$ is the main step required for proving that T is a functor.

REMARK 3.20. If M is a manifold of class C^k for some finite $k \in \mathbb{N}$, then the definition of tangent spaces requires a slight adjustment since the notion of *smooth* paths in M might not make sense; it is good enough however (and gives an equivalent definition) if we consider all paths

$\gamma : (-\epsilon, \epsilon) \to M$ of class C^1. Inspecting the proof of Corollary 3.12 now reveals that TM is naturally a manifold of class C^{k-1}; one derivative is lost because the transition maps for TM involve derivatives of the transition maps for M. Similarly, if $f : M \to N$ is of class C^r with $1 \leqslant r \leqslant k$, then the tangent map $Tf : TM \to TN$ can be defined as a map of class C^{r-1}.

4. Submanifolds

The overarching message of this lecture will be that sometimes, understanding what is happening in a manifold is just a matter of finding the right coordinates.

4.1. Partial derivatives and differentials. There are two special situations in which the tangent map of $f : M \to N$ can be expressed in slightly more convenient forms. First, if $\mathcal{U} \subset \mathbb{R}^n$ is an open subset of Euclidean space, M is a manifold and $f : \mathcal{U} \to M$ is smooth, then f can be regarded (without needing to make a choice of coordinates) as an M-valued function of n variables, $f(x^1, \dots, x^n)$. For each point $x_0 = (x_0^1, \dots, x_0^n) \in \mathcal{U}$, f now determines n smooth paths through $f(x_0)$, namely

$$\gamma_j(t) := f(x_0^1, \dots, x_0^{j-1}, x_0^j + t, x_0^{j+1}, \dots, x_0^n), \qquad j = 1, \dots, n.$$

The equivalence classes of these paths are called the **partial derivatives** of f at x_0,

$$\partial_j f(x_0) := \frac{\partial f}{\partial x^j}(x_0) := [\gamma_j] \in T_{f(x_0)}M.$$

They are actually just particular values of the tangent map, i.e. $\partial_j f(x_0) = T_{x_0} f(e_j)$, where we are using the fact that $T_{x_0}\mathcal{U}$ is canonically isomorphic to $\mathbb{R}^n$ (see Example 3.7) and thus comes with a canonical basis $e_1, \dots, e_n$. The n tangent vectors $\partial_1 f(x_0), \dots, \partial_n f(x_0) \in T_{f(x_0)}M$ all together thus contain the same information as the tangent map $T_{x_0} f : T_{x_0}\mathcal{U} \to T_{f(x_0)}M$.

The second special situation is in some dense dual to the first: we consider a smooth function on a smooth manifold M with values in a finite-dimensional vector space V,

$$f : M \to V.$$

The most important special case of this is when $V = \mathbb{R}$, so that f is a real-valued function. Taking advantage again of the canonical isomorphisms $T_{f(p)}V = V$ from Example 3.7, we can rewrite $Tf(X) \in T_{f(p)}V$ for each $p \in M$ and $X \in T_pM$ as a vector in V, denoted by $df(X) \in V$. This associates to every smooth function $f : M \to V$ a smooth function

$$df : TM \to V,$$

called the **differential** (*Differential*) of f. We will denote its restriction to each individual tangent space T_pM for $p \in M$ by

$$d_p f : T_pM \to V.$$

In terms of equivalence classes of paths through p, a direct formula for $d_p f$ is given by

$$d_p f([\gamma]) = (f \circ \gamma)'(0), \tag{4.1}$$

and one can deduce from Lemma 3.15 that this is independent of the choice of path γ in the equivalence class, and moreover, $d_p f : T_pM \to V$ is a linear map. In particular, for a smooth real-valued function $f : M \to \mathbb{R}$, $d_p f$ is an element of the cotangent space at p,

$$d_p f \in T_p^* M \qquad (\text{for } f : M \to \mathbb{R}).$$

This makes the differentials df of smooth real-valued functions $f : M \to \mathbb{R}$ into our first examples of *differential forms*; we will have a lot more to say about them when we discuss integration in a few weeks.

EXAMPLE 4.1. The differentials defined above directly generalize the linear map $d_p x : T_p M \to \mathbb{R}^n$ in (3.2), which can be associated to any smooth chart $(\mathcal{U}, x)$ on M and a point $p \in \mathcal{U}$. This map can also be constructed out of the differentials of the coordinate functions $x^1, \ldots, x^n : \mathcal{U} \to \mathbb{R}$; it is given by

$$d_p x(X) = (d_p x^1(X), \ldots, d_p x^n(X)) \in \mathbb{R}^n.$$

4.2. The inverse function theorem. In the examples of manifolds we have dealt with so far, we have always had charts that were explicitly constructed, but such explicit constructions are not always convenient in more general situations. A nice tool for obtaining less explicit but often more useful constructions of charts is provided by the inverse function theorem from first-year analysis. Let us recall the statement:

THEOREM (inverse function theorem). *Suppose $\mathcal{U} \subset \mathbb{R}^n$ is open, $f : \mathcal{U} \to \mathbb{R}^n$ is a map of class C^k for some $k \in \mathbb{N} \cup \{\infty\}$, and $x_0 \in \mathcal{U}$ is a point at which the derivative $Df(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism. Then there exist open neighborhoods $x_0 \in \Omega \subset \mathcal{U}$ and $f(x_0) \in \Omega' \subset \mathbb{R}^n$ such that f maps Ω bijectively onto Ω' and the inverse $(f|_\Omega)^{-1} : \Omega' \to \Omega$ is also of class C^k.* □

We will now turn this standard analytical result into a pair of criteria for proving that certain maps we construct define smooth charts.

LEMMA 4.2. *Suppose M is a smooth n-manifold, $\mathcal{U} \subset \mathbb{R}^n$ is an open set, $\varphi : \mathcal{U} \to M$ is a smooth map and $x_0 \in \mathcal{U}$ is a point at which the partial derivatives $\partial_1\varphi(x_0), \ldots, \partial_n\varphi(x_0)$ form a basis of $T_{\varphi(x_0)}M$. Then there exist open neighborhoods $x_0 \in \Omega \subset \mathcal{U}$ and $p := \varphi(x_0) \in \mathcal{O} \subset M$ such that φ maps Ω bijectively onto $\mathcal{O}$ and $(\mathcal{O}, (\varphi|_\Omega)^{-1})$ defines a smooth chart on M.*

PROOF. Choose any smooth chart $(\mathcal{V}, y)$ on M with $p = \varphi(x_0) \in \mathcal{V}$, and observe that $d_p y(\partial_j \varphi(x_0)) = \partial_j (y \circ \varphi)(x_0)$In for each $j = 1, \ldots, n$. Since $d_p y : T_p M \to \mathbb{R}^n$ is an isomorphism, our assumption on the basis $\partial_1\varphi(x_0), \ldots, \partial_n\varphi(x_0) \in T_p M$ means that $\partial_1(y \circ \varphi)(x_0), \ldots, \partial_n(y \circ \varphi)(x_0)$ is similarly a basis of $\mathbb{R}^n$, which is equivalent to saying that the linear map $D(y \circ \varphi)(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism. The inverse function theorem thus provides open neighborhoods $x_0 \in \Omega \subset \mathcal{U}$ and $y(p) \in \Omega' \subset \mathbb{R}^n$ such that $y \circ \varphi$ is a diffeomorphism between Ω and Ω', implying that $\varphi = y^{-1} \circ (y \circ \varphi)$ sends Ω bijectively to an open neighborhood $\mathcal{O} := y^{-1}(\Omega')$ of p. Denoting the inverse of this bijection by $x : \mathcal{O} \to \Omega \subset \mathbb{R}^n$, the transition map $y \circ x^{-1}$ is now just $y \circ \varphi|_\Omega$, so it is smooth and has a smooth inverse. □

LEMMA 4.3. *Suppose M is a smooth n-manifold, $\mathcal{U} \subset M$ is an open set, $x^1, \ldots, x^n : \mathcal{U} \to \mathbb{R}$ are smooth functions and $p \in \mathcal{U}$ is a point such that the differentials $d_p x^1, \ldots, d_p x^n$ form a basis of $T_p^* M$. Then there exists an open neighborhood $p \in \mathcal{O} \subset \mathcal{U}$ such that $(\mathcal{O}, x)$ with $x := (x^1, \ldots, x^n) : \mathcal{O} \to \mathbb{R}^n$ defines a smooth chart on M.*

PROOF. Since $d_p x^1, \ldots, d_p x^n$ is a basis of $T_p^* M$, it is dual to a unique basis $X_1, \ldots, X_n$ of $T_p M$, meaning the two bases are related by

$$d_p x^i(X_j) = \delta^i_j := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Define the linear map $d_p x := (d_p x^1, \ldots, d_p x^n) : T_p M \to \mathbb{R}^n$ as in Example 4.1, so $d_p x$ is the tangent map $T_p x : T_p M \to T_{x(p)}\mathbb{R}^n$ after identifying $T_{x(p)}\mathbb{R}^n = \mathbb{R}^n$. Since $d_p x$ sends the basis $X_1, \ldots, X_n$ to the standard basis of $\mathbb{R}^n$, it is an isomorphism. Now if $(\mathcal{V}, y)$ is any smooth chart with $p \in \mathcal{V}$, the map $x \circ y^{-1}$ is smooth on a neighborhood of p, and the chain rule gives

$$D(x \circ y^{-1})(y(p)) = d_p x \circ (d_p y)^{-1},$$

hence the latter is also an isomorphism $\mathbb{R}^n \to \mathbb{R}^n$. The inverse function theorem now provides open neighborhoods $y(p) \in \Omega \subset \mathbb{R}^n$ and $x(p) \in \Omega' \subset \mathbb{R}^n$ such that $x \circ y^{-1}$ is a diffeomorphism from

Ω onto Ω', so $\mathcal{O} := y^{-1}(\Omega) = x^{-1}(\Omega')$ is then a neighborhood of p on which the restriction of x defines a chart that is smoothly compatible with $(\mathcal{V}, y)$. □

4.3. Slice charts. We have used the word "submanifold" already a few times in an informal way, e.g. the unit circle S^1 is a manifold that lives inside the manifold $\mathbb{R}^2$, so we called it a submanifold. It is now time to clarify more precisely what this word means.

The archetypal example of a submanifold is a linear subspace of a vector space, for instance

$$\mathbb{R}^\ell \times \{0\} = \left\{(x^1, \dots, x^\ell, 0, \dots, 0) \in \mathbb{R}^n \;\middle|\; (x^1, \dots, x^\ell) \in \mathbb{R}^\ell\right\} \subset \mathbb{R}^n.$$

Basic results in linear algebra imply that *any* ℓ-dimensional subspace of an n-dimensional vector space looks like this example after a suitable linear change of coordinates. The notion of a smooth submanifold generalizes this by allowing nonlinear (but smooth) changes of coordinates.

DEFINITION 4.4. A chart $(\mathcal{U}, x)$ on an n-manifold M is called an ℓ-dimensional **slice chart** (*Bügelkarte*) for a subset $L \subset M$ if

$$L \cap \mathcal{U} = x^{-1}(\mathbb{R}^\ell \times \{0\}),$$

i.e. the points in $\mathcal{U}$ belong to L if and only if their coordinates $x^{\ell+1}, \dots, x^n$ vanish.

DEFINITION 4.5. Suppose M is a smooth n-manifold. A subset $L \subset M$ is called an ℓ-**dimensional smooth submanifold** (*Untermannigfaltigkeit*) of M if M admits a collection of smooth slice charts for L whose domains cover every point of L.

REMARK 4.6. More generally, if M is a manifold of class C^k but not necessarily smooth, one can speak of *submanifolds of class* C^k, in which the transition maps between slice charts are required to be of class C^k. Note that a C^k-manifold can also be regarded as a C^r-manifold for any $r \leqslant k$, so under this condition it makes sense to talk about C^r-submanifolds, but e.g. there is no such thing as a *smooth* submanifold of M if the latter is of class C^k for some $k < \infty$ but not equipped with a smooth structure.

EXAMPLE 4.7. The smooth structure we constructed on $S^1 \subset \mathbb{R}^2$ in Lecture 1 was obtained from polar coordinates by restricting to the unit circle $\{r = 1\}$; this gave rise to two charts $(\mathcal{U}, \theta)$ and $(\mathcal{V}, \phi)$, where θ and ϕ both had the meaning of an angle in polar coordinates, but with different ranges of values, namely $\theta(\mathcal{U}) = (0, 2\pi)$ and $\phi(\mathcal{V}) := (-\pi, \pi)$. These two coordinates were defined on open subsets of S^1, but they also have natural extensions to open subsets of $\mathbb{R}^2$, namely

$$\mathcal{U}' := \left\{tv \in \mathbb{R}^2 \;\middle|\; v \in \mathcal{U},\ t > 0\right\}, \qquad \mathcal{V}' := \left\{tv \in \mathbb{R}^2 \;\middle|\; v \in \mathcal{V},\ t > 0\right\}.$$

The radial coordinate r is defined on $\mathbb{R}^2 \backslash \{0\}$ and takes all positive values; if we now set $\rho := r - 1$ so that $\{r = 1\} = \{\rho = 0\}$, we obtain a pair of smoothly compatible slice charts $(\mathcal{U}', (\theta, \rho))$ and $(\mathcal{V}', (\phi, \rho))$ for S^1 such that $S^1 \subset \mathcal{U}' \cup \mathcal{V}'$. This means that S^1 is a smooth submanifold of $\mathbb{R}^2$.

One can similarly turn the atlas for S^2 in Exercise 1.7 into a family of slice charts to prove that S^2 is a submanifold of $\mathbb{R}^3$. In practice, however, constructing slice charts by hand is not usually necessary, as we will see in §4.4 that some much more general and powerful tools for this purpose are provided by the inverse function theorem.

Let us first clarify the fact that a submanifold of a manifold is also a manifold in its own right.

PROPOSITION 4.8. *If L is an ℓ-dimensional C^k-submanifold of an n-dimensional C^k-manifold M, then L inherits naturally from M the structure of an ℓ-dimensional C^k-manifold such that the inclusion map $L \hookrightarrow M$ is of class C^k. Moreover, for each $p \in L$, the tangent space T_pL is naturally an ℓ-dimensional linear subspace of T_pM.*

PROOF. We associate to every slice chart $(\mathcal{U}, x)$ for $L \subset M$ a chart of the form $(\mathcal{U} \cap L, x_L)$ on L, where we use the coordinate projection $\pi_\ell(x^1, \ldots, x^n) := (x^1, \ldots, x^\ell)$ to define

$$x_L = \pi_\ell \circ x|_{\mathcal{U} \cap L} : \mathcal{U} \cap L \to \mathbb{R}^\ell.$$

By assumption, L can be covered by slice charts, so the collection of all charts of this form defines an atlas on L. Given two such charts $(\mathcal{U} \cap L, x_L)$ and $(\mathcal{V} \cap L, y_L)$ derived from two C^k-compatible slice charts $(x, \mathcal{U})$ and $(y, \mathcal{V})$, the transition map $y \circ x^{-1}$ preserves the subspace $\mathbb{R}^\ell \times \{0\} \subset \mathbb{R}^n$, and its restriction to the intersection of its domain with this subspace is the transition map $y_L \circ x_L^{-1}$, which is therefore of class C^k. Moreover, the fact that M is metrizable and separable implies the same for L by Exercise 2.17, thus L is a C^k-manifold. The local coordinate expression for the inclusion $i : L \hookrightarrow M$ with respect to any slice chart $(\mathcal{U}, x)$ and the associated chart $(\mathcal{U} \cap L, x_L)$ on L is $(x^1, \ldots, x^\ell) \mapsto (x^1, \ldots, x^\ell, 0, \ldots, 0)$, which is clearly smooth, thus the inclusion is of class C^k.[22] For each $p \in L$, the tangent map $T_p i : T_p L \to T_p M$ is simply the canonical inclusion $T_p L \hookrightarrow T_p M$ defined by regarding each path in L as a path in M. Since its image is a linear subspace, it gives a canonical isomorphism of $T_p L$ to a linear subspce of $T_p M$. □

Whenever we speak of a submanifold $L \subset M$ from now on, we will assume that L is endowed with the differentiable structure described in Proposition 4.8, so that it can also be regarded as a manifold in its own right. We will often make use of the canonical identification of tangent spaces $T_p L$ with subspaces of $T_p M$, especially in the case $M = \mathbb{R}^n$, where (in light of Example 3.7) this identification allows us to view each tangent space $T_p L$ as a subspace of $\mathbb{R}^n$.

EXERCISE 4.9. Assume in the following that M and N are both C^k-manifolds and $f : M \to N$ is a map of class C^k. Prove:

(a) For any C^k-submanifold $L \subset M$, the restriction $f|_L : L \to N$ is also a map of class C^k.
(b) If $L \subset N$ is a C^k-submanifold such that $f(M) \subset L$, then the resulting map $f : M \to L$ is also of class C^k.

4.4. Immersions and submersions.

DEFINITION 4.10. A smooth map $f : M \to N$ is called an **immersion** at $p \in M$ if the linear map $T_p f : T_p M \to T_{f(p)} N$ is injective, and similarly, f is a **submersion** at p if $T_p f : T_p M \to T_{f(p)} N$ is surjective. If one says that f is an immersion/submersion without specifying a point p, the meaning is that it is true for *all* points in M. One sometimes uses the notation

$$f : M \looparrowright N$$

to indicate when f is an immersion.

Recall that for any two finite-dimensional vector spaces V, W, the sets of linear maps $V \to W$ that are injective or surjective are open. It follows that if f is an immersion or submersion at some point $p \in M$, then this is also true on a *neighborhood* of p; equivalently, the set of points at which f is an immersion or submersion is open.

There is a good reason to single out these two particular classes of smooth maps between manifolds: it turns out that up to choices of smooth coordinates near $p \in M$ and $f(p) \in N$, all immersions look the same, and similarly for all submersions. This fact will give us a new user-friendly tool for identifying smooth submanifolds. The main tool required in its proof is the inverse function theorem, or more precisely, the two lemmas in §4.2 that used the inverse function theorem to construct charts.

[22] Recall that if both L and M are manifolds of class C^k but $k < \infty$, then it does not make sense to say that the inclusion $L \hookrightarrow M$ is smooth, even though it looks smooth in the particular local coordinates we chose. The point is that one could also choose different coordinates in which it would still appear to be a map of class C^k, but not necessarily C^∞.

Theorem 4.11. *Assume M is a smooth m-manifold, N is a smooth n-manifold, $f : M \to N$ is a smooth map, $p \in M$ and $q = f(p) \in N$. If f is either an immersion or a submersion at p, then there exist smooth charts $(\mathcal{U}, x)$ on M with $x(p) = 0 \in \mathbb{R}^m$ and $(\mathcal{V}, y)$ on N with $y(q) = 0 \in \mathbb{R}^n$ such that the coordinate expression $y \circ f \circ x^{-1}$ for f is given by*

$$\mathbb{R}^m \ni (x^1, \dots, x^m) \mapsto \begin{cases} (x^1, \dots, x^n) \in \mathbb{R}^n & \text{if } m \geqslant n \text{ (submersion case)}, \\ (x^1, \dots, x^m, 0, \dots, 0) \in \mathbb{R}^n & \text{if } m < n \text{ (immersion case)}. \end{cases}$$

Proof. Assume first that $T_p f : T_p M \to T_{f(p)} N$ is injective, so $n \geqslant m$, and set $\ell := n - m$. Choose a smooth chart $(\mathcal{U}, x)$ on M with $p \in \mathcal{U}$ and $x(p) = 0 \in \mathbb{R}^m$; note that the latter can be assumed without loss of generality by taking any chart with $p \in \mathcal{U}$ and composing the map $\mathcal{U} \to \mathbb{R}^n$ with a translation on $\mathbb{R}^n$ sending the image of p to the origin. With this understood, $\Omega := x(\mathcal{U}) \subset \mathbb{R}^m$ is an open neighborhood of the origin, and we observe that $F := f \circ x^{-1} : \Omega \to N$ is now a smooth map such that $F(0) = q$ and $T_0 F = T_p f \circ (d_p x)^{-1} : \mathbb{R}^m \to T_q N$ is injective. The latter is equivalent to the condition that the partial derivatives $\partial_1 F(0), \dots, \partial_m F(0) \in T_q N$ are linearly independent.

We claim that after possibly shrinking Ω to a smaller neighborhood of $0 \in \mathbb{R}^m$, and choosing $\epsilon > 0$ sufficiently small, $F : \Omega \to N$ can be extended to a smooth map

$$\widetilde{F} : \Omega \times (-\epsilon, \epsilon)^\ell \to N$$

such that $\widetilde{F}(x^1, \dots, x^m, 0, \dots, 0) = F(x^1, \dots, x^m)$ and the partial derivatives $\partial_1 \widetilde{F}, \dots, \partial_n \widetilde{F}$ at the origin form a basis of $T_q N$. This extension is not canonical, but it is also not difficult: if N were simply $\mathbb{R}^n$, we could define it by choosing any extension of the linearly independent set $\partial_1 F(0), \dots, \partial_m F(0)$ to a basis $\partial_1 F(0), \dots, \partial_m F(0), Y_{m+1}, \dots, Y_n$ of $T_q N$ and then defining

$$\widetilde{F}(x^1, \dots, x^n) := F(x^1, \dots, x^m) + \sum_{j=m+1}^{n} x^j Y_j.$$

This formula does not make sense in general if N is not a vector space, but one could more generally choose a chart on N near q in order to express F in local coordinates, and define the extension in this way in coordinates. Lemma 4.2 now implies that on a sufficiently small neighborhood of $0 \in \mathbb{R}^n$, $\widetilde{F}$ can be inverted to define a chart $(\mathcal{V}, y)$ on N with the stated properties.

Next suppose $T_p f : T_p M \to T_{f(p)} N$ is surjective, thus $m \geqslant n$, and we can set $\ell := m - n$. The idea now is to choose any chart $(\mathcal{V}, y)$ on N with $y(q) = 0$ and define the first n coordinates over the neighborhood $f^{-1}(\mathcal{V}) \subset M$ of p by

$$x^i := y^i \circ f, \qquad i = 1, \dots, n.$$

Writing $\hat{x} := (x^1, \dots, x^n) : f^{-1}(\mathcal{V}) \to \mathbb{R}^n$, we have $d_p \hat{x} = d_q y \circ T_p f$, thus $d_p \hat{x} : T_p M \to \mathbb{R}^n$ is surjective, which is equivalent to the condition that the n covectors $d_p x^1, \dots, d_p x^n \in T_p^* M$ are linearly independent.

To define the remaining ℓ coordinates on M near p, first choose an extension of the linearly-independent set $d_p x^1, \dots, d_p x^n$ to a basis $d_p x^1, \dots, d_p x^n, \Lambda^{n+1}, \dots, \Lambda^m$ of $T_p^* M$. For each $i = n+1, \dots, m$, we can then define a smooth function x^i on a neighborhood of p such that $x^i(p) = 0$ and $d_p x^i = \Lambda^i$; this is another step that would be trivial to carry out if M were the vector space $\mathbb{R}^m$, so the idea is to choose a chart near p and write down suitable functions in local coordinates. With this done, Lemma 4.3 implies that after possibly shrinking to a smaller neighborhood $\mathcal{U} \subset M$ of p, $x = (x^1, \dots, x^m)$ becomes a smooth chart with the desired properties. □

Remark 4.12. For a continuous map $f : M \to N$ between topological manifolds, one can define f to be a *topological immersion* or *topological submersion* at $p \in M$ if there exist continuous charts near p and $q := f(p)$ in which f satisfies the coordinate formula in Theorem 4.11. Note that without having at least one continuous derivative at our disposal, there is no alternative way to characterize either of these conditions in terms of a tangent map being injective or surjective, nor is there any inverse function theorem available for proving such statements. On the other hand, Theorem 4.11 does make sense in the setting of C^k-manifolds for any $k \in \mathbb{N}$; in this case one must assume that $f : M \to N$ is of class C^k, and the resulting charts will be as well. (One should not be fooled by the fact that f will then *look* like a smooth map with respect to those charts—if $k < \infty$, it will not look smooth after arbitrary changes of C^k-coordinates.)

4.5. Embeddings and regular level sets. We now have enough technology to produce many more examples of submanifolds.

Definition 4.13. A smooth map $f : M \to N$ is called an **embedding** (*Einbettung*) if it is an injective immersion whose inverse $f(M) \stackrel{f^{-1}}{\to} M$ is also continuous. The notation

$$f : M \hookrightarrow N$$

is sometimes used to indicate that f is an embedding.

The typical example of an embedding is the natural inclusion $M \hookrightarrow N$ that exists whenever M is a submanifold of N. The next result states that, up to diffeomorphism, all examples are this one.

Theorem 4.14. *If $f : M \to N$ is an embedding, then its image $f(M)$ is a smooth submanifold of N.*

Proof. Suppose $q \in f(M)$. By injectivity, there is a unique point $p \in M$ such that $f(p) = q$, and Theorem 4.11 provides charts $(\mathcal{U}, x)$ on M and $(\mathcal{V}, y)$ on N with $x(p) = 0$ and $y(q) = 0$ such that $y \circ f \circ x^{-1}$ takes the form $(x^1, \dots, x^m) \mapsto (x^1, \dots, x^m, 0, \dots, 0)$. Since the inverse $f(M) \to M$ is also continuous, we are free to assume after possibly shrinking $\mathcal{V} \subset N$ to a smaller neighborhood of q that

$$f^{-1}(\mathcal{V} \cap f(M)) \subset \mathcal{U},$$

or in other words, $\mathcal{V} \cap f(M) = f(\mathcal{U})$. This proves that $(\mathcal{V}, y)$ is a slice chart for the subset $f(M)$. □

The following consequence appears in some books as an alternative definition of the notion of a submanifold:

Corollary 4.15. *A subset $L \subset M$ of a smooth manifold M is a smooth submanifold if and only if it admits a smooth structure for which the inclusion map $L \hookrightarrow M$ is a smooth embedding.* □

It is worth pausing a moment to consider what an immersion $f : M \looparrowright N$ can look like if it is *not* an embedding. Theorem 4.11 implies that every immersion is *locally* an embedding, i.e. for every $p \in M$, one can find a neighborhood $\mathcal{U} \subset M$ of p such that $f|_{\mathcal{U}} : \mathcal{U} \hookrightarrow N$ is an embedding and $f(\mathcal{U}) \subset N$ is therefore a submanifold. On the other hand, f may fail to be an embedding *globally* because it is not injective, meaning it has self-intersections $f(p) = f(p')$ with $p \neq p'$. The notation "$f : M \looparrowright N$" is meant to evoke this possibility by allowing the arrow to loop around and intersect itself. A classic example of a non-injective immersion is the picture of the Klein bottle in Figure 6, which shows the image of an immersion of a compact smooth 2-manifold into $\mathbb{R}^3$. Images of immersions are sometimes called **immersed submanifolds** in the literature, though I am personally not fond of this terminology,[23] so I will not use it.

[23] I have two objections to the term "immersed submanifold": first, it sounds as if it should be a type of submanifold, but it isn't. Second, one cannot always uniquely recover the manifold M from the image of an immersion $M \looparrowright N$. For example (the following is only for readers with a background in topology), a closed surface Σ_g of genus $g \geqslant 2$ admits smooth covering maps $\Sigma_h \to \Sigma_g$ by surfaces of arbitrarily large genus h (the degree of the cover will be correspondingly large). If one chooses an embedding of Σ_g into $\mathbb{R}^3$, one obtains a submanifold that is also the image of an immersion $\Sigma_h \looparrowright \mathbb{R}^3$ for arbitrarily large values of h.

For slightly subtler reasons, an injective immersion can also fail to be an embedding:

EXAMPLE 4.16. Let $N = \mathbb{R}^2$ and $M = \mathbb{R} \amalg (0, \pi)$, and defne the immersion $f : M \looparrowright \mathbb{R}^2$ by

$$f(t) := (t, 0) \qquad \text{for } t \in \mathbb{R},$$
$$f(\theta) := (\cos\theta, \sin\theta) \qquad \text{for } \theta \in (0, \pi).$$

Omitting the points 0 and π from the interval $(0, \pi)$ makes this map an injective immersion, but the inverse $f(M) \xrightarrow{f^{-1}} M$ is discontinuous at the two points $(\pm 1, 0)$, which are precisely the points at which it fails to be a submanifold.

Turning our attention to submersions, we can now state a popular corollary of the implicit function theorem that you may have heard referred to before as the "regular value theorem".

DEFINITION 4.17. For a smooth map $f : M \to N$, $p \in M$ is called a **regular point** (*regulärer Wert*) of f if f is a submersion at p, and a **critical point** (*kritischer Wert*) otherwise. A point $q \in N$ is a **critical value** (*kritischer Wert*) of f if $q = f(p)$ for some critical point p, and q is otherwise called a **regular value** (*regulärer Wert*) of f.

THEOREM 4.18 (implicit function theorem). *For any smooth map $f : M \to N$ with regular value $q \in N$, $L := f^{-1}(q) \subset N$ is a smooth submanifold with $\dim L = \dim M - \dim N$, and its tangent space at any point $p \in L$ is $T_pL = \ker T_pf \subset T_pM$.*

PROOF. For each $p \in L = f^{-1}(q)$, f is by assumption a submersion at p, so Theorem 4.11 provides charts x near p and y near q such that $x(p)$ and $y(q)$ are both the origin in their respective Euclidean spaces and $y \circ f \circ x^{-1}$ becomes the map $(x^1, \ldots, x^m) \mapsto (x^1, \ldots, x^n)$. The zero-set of this map is a neighborhood of p in $f^{-1}(q)$ as seen in the x-coordinates, thus x is a slice chart. To see that $T_pL = \ker T_pf$, observe first that for any path γ in L through p, $f \circ \gamma$ is a constant path at $q \in N$, thus $T_pf([\gamma]) = 0 \in T_qN$, proving $T_pL \subset \ker T_pf$. The rest is dimension counting, as the surjectivity of $T_pf : T_pM \to T_qN$ implies

$$\dim T_pL = \dim L = \dim M - \dim N = \dim T_pM - \dim T_qN = \dim \ker T_pf.$$

□

Submanifolds of the form $f^{-1}(q) \subset M$ for regular values $q \in N$ are sometimes called **regular level sets** of f. In particular, a submersion $f : M \to N$ is distinguished by the property that *all* of its level sets are regular, and are thus smooth submanifolds.

4.6. Examples. We now have a *very* easy way of proving that simple examples like the unit spheres $S^n \subset \mathbb{R}^{n+1}$ really are smooth submanifolds.

EXAMPLE 4.19. Define $f : \mathbb{R}^{n+1} \to \mathbb{R}$ in terms of the standard Euclidean inner product by $f(x) = |x|^2 = \langle x, x\rangle$. This is a smooth map, with differential at any point $x \in \mathbb{R}^{n+1}$ given by $d_xf(v) = 2\langle x, v\rangle$, so it is a submersion everywhere except at the origin. This makes $S^n = f^{-1}(1)$ into a smooth submanifold of dimension $(n + 1) - 1 = n$, so in particular, S^n inherits a natural smooth structure for which the inclusion $S^n \hookrightarrow \mathbb{R}^{n+1}$ is a smooth embedding. The kernel of d_xf at a point $x \in S^n$ is the orthogonal complement of x, hence

$$T_xS^n = x^\perp \subset \mathbb{R}^{n+1}.$$

EXAMPLE 4.20. The smooth map $f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto xy$ has only one critical point, at $(x, y) = (0, 0)$, thus $f^{-1}(t)$ is a smooth submanifold (a hyperbola) for every $t \neq 0$, and so is $f^{-1}(0)\backslash\{(0, 0)\}$, but $f^{-1}(0)$ fails to be a submanifold at the origin.

EXERCISE 4.21. Identifying the torus $\mathbb{T}^2$ with $\mathbb{R}^2/\mathbb{Z}^2$ via Exercise 3.4, find an explicit formula for an embedding $\mathbb{T}^2 \hookrightarrow \mathbb{R}^3$ whose image looks like Figure 5.

For the next set of exercises, the symbol $\mathbb{F}$ always denotes either the real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$, and we denote the vector space of m-by-n matrices over $\mathbb{F}$ by

$$\mathbb{F}^{m\times n} := \{m\text{-by-}n \text{ matrices over } \mathbb{F}\}.$$

If $\mathbb{F} = \mathbb{R}$, this is a real vector space of dimension mn. In the case $\mathbb{F} = \mathbb{C}$, it is a complex vector space of this same dimension, which means it can also be regarded as a *real* vector space of dimension $2mn$. (Indeed, if V is any complex vector space with complex basis $v_1, \ldots, v_k$, then a basis of V as a *real* vector space is given by $v_1, iv_1, \ldots, v_k, iv_k$.) Since they are vector spaces, $\mathbb{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ carry natural smooth structures and are thus smooth manifolds of dimensions mn and $2mn$ respectively. For $m = n$, there is a distinguished open subset

$$\mathrm{GL}(n, \mathbb{F}) = \left\{\mathbf{A} \in \mathbb{F}^{n\times n} \bigm| \mathbf{A} \text{ is invertible}\right\},$$

which is therefore also naturally a smooth manifold of dimension n^2 or (in the complex case) $2n^2$. That $\mathrm{GL}(n, \mathbb{F}) \subset \mathbb{F}^{n\times n}$ is open can be deduced easily from the observation that the determinant

$$\det : \mathbb{F}^{n\times n} \to \mathbb{F}$$

defines a continuous function for which $\mathrm{GL}(n, \mathbb{F}) = \det^{-1}(\mathbb{F}\backslash\{0\})$. In fact, $\det(\mathbf{A})$ is a polynomial in the entries of $\mathbf{A}$, which are all linear functions of $\mathbf{A}$, thus $\det : \mathbb{F}^{n\times n} \to \mathbb{F}$ is a smooth real- or complex-valued function. By Cramer's rule, the function

$$\mathrm{GL}(n, \mathbb{F}) \to \mathrm{GL}(n, \mathbb{F}) : \mathbf{A} \mapsto \mathbf{A}^{-1}$$

is also smooth.

EXERCISE 4.22. The n-dimensional **orthogonal group** $\mathrm{O}(n) \subset \mathbb{R}^{n\times n}$ is the set of all real n-by-n matrices $\mathbf{A}$ with the property

$$\mathbf{A}^T\mathbf{A} = \mathbb{1},$$

where $\mathbb{1}$ is the n-by-n identity matrix and $\mathbf{A}^T$ denotes the *transpose* of $\mathbf{A}$, i.e. if $\mathbf{A}$ has entries A_{ij}, then the corresponding entries of $\mathbf{A}^T$ are A_{ji}. This is precisely the set of all linear transformations $\mathbb{R}^n \to \mathbb{R}^n$ which preserve the Euclidean inner product, which means geometrically that they preserve lengths of vectors and angles between them. We will show in this exercise that $\mathrm{O}(n)$ is a smooth submanifold of $\mathbb{R}^{n\times n}$.

(a) Define the linear subspace consisting of all symmetric matrices,

$$\Sigma(n) := \left\{\mathbf{A} \in \mathbb{R}^{n\times n} \bigm| \mathbf{A} = \mathbf{A}^T\right\} \subset \mathbb{R}^{n\times n}.$$

There is a map

$$f : \mathbb{R}^{n\times n} \to \Sigma(n) : \mathbf{A} \mapsto \mathbf{A}^T\mathbf{A},$$

such that the orthogonal group is the level set $\mathrm{O}(n) = f^{-1}(\mathbb{1})$. The entries of $f(\mathbf{A})$ are quadratic functions of the entries of $\mathbf{A}$, thus f is clearly a smooth map. Show that its derivative at any $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the linear map

$$Df(\mathbf{A}) : \mathbb{R}^{n\times n} \to \Sigma(n) : \mathbf{H} \mapsto \mathbf{A}^T\mathbf{H} + \mathbf{H}^T\mathbf{A}.$$

Hint: In theory you can do this by computing all the partial derivatives of f with respect to the entries of $\mathbf{A}$, but it's much, much *easier to use the definition of the derivative, i.e. regarding $\mathbb{R}^{n\times n}$ and $\Sigma(n)$ simply as vector spaces, show that a "remainder" formula of the form*

$$f(\mathbf{A} + \mathbf{H}) = f(\mathbf{A}) + Df(\mathbf{A})\mathbf{H} + R(\mathbf{H}) \cdot |\mathbf{H}|$$

with $\lim_{\mathbf{H}\to 0} R(\mathbf{H}) = 0$ is satisfied. One useful thing you may want to assume: for a reasonable choice of norm on $\mathbb{R}^{n\times n}$, matrix products satisfy $|\mathbf{AB}| \leqslant |\mathbf{A}||\mathbf{B}|$.

(b) Show that $Df(\mathbf{A})$ is surjective if $\mathbf{A} \in \mathrm{O}(n)$. In fact, you won't even need to assume $\mathbf{A} \in \mathrm{O}(n)$, but it *is* useful to assume that $\mathbf{A}$ is invertible (which is automatically true for orthogonal matrices). It is also *crucial* that the target space is $\Sigma(n)$ rather than the entirety of $\mathbb{R}^{n \times n}$—$Df(\mathbf{A})$ is certainly *not* surjective onto $\mathbb{R}^{n \times n}$.

(c) It follows now from the implicit function theorem that $\mathrm{O}(n)$ is a smooth submanifold of $\mathbb{R}^{n \times n}$. What is its dimension? (For a sanity check I will tell you: $\dim \mathrm{O}(2) = 1$ and $\dim \mathrm{O}(3) = 3$.)

(d) Show that $T_{\mathbb{1}} \mathrm{O}(n) \subset T_{\mathbb{1}} \mathbb{R}^{n \times n} = \mathbb{R}^{n \times n}$ is the space of all *antisymmetric* matrices $\mathbf{H}$, i.e. those which satisfy $\mathbf{H}^T = -\mathbf{H}$.

Exercise 4.23. The complex analogue of Exercise 4.22 involves the **unitary group**

$$\mathrm{U}(n) = \left\{ \mathbf{A} \in \mathbb{C}^{n \times n} \;\middle|\; \mathbf{A}^\dagger \mathbf{A} = \mathbb{1} \right\},$$

where $\mathbf{A}^\dagger$ denotes the Hermitian adjoint of $\mathbf{A}$, defined as the complex conjugate of its transpose. Prove that $\mathrm{U}(n)$ is a smooth submanifold of $\mathbb{C}^{n \times n}$, compute its dimension, and show

$$T_{\mathbb{1}} \mathrm{U}(n) = \left\{ \mathbf{H} \in \mathbb{C}^{n \times n} \;\middle|\; \mathbf{H}^\dagger = -\mathbf{H} \right\}.$$

Exercise 4.24. The **special linear group** over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is defined by

$$\mathrm{SL}(n, \mathbb{F}) = \left\{ \mathbf{A} \in \mathbb{F}^{n \times n} \;\middle|\; \det(\mathbf{A}) = 1 \right\}.$$

(a) Show that the derivative of $\det : \mathbb{F}^{n \times n} \to \mathbb{F}$ at $\mathbb{1}$ is given by the **trace** (*Spur*):

$$D(\det)(\mathbb{1})\mathbf{H} = \mathrm{tr}(\mathbf{H}).$$

Hint: Write $\mathbf{H}$ *in terms of* n *column vectors as* $\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$, *so*

$$\det(\mathbb{1} + t\mathbf{H}) = \det \begin{pmatrix} \mathbf{e}_1 + t\mathbf{v}_1 & \cdots & \mathbf{e}_n + t\mathbf{v}_n \end{pmatrix},$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *denotes the standard basis of* $\mathbb{F}^n$. *Differentiate this expression with respect to* t *at* $t = 0$, *using the fact that the determinant of a matrix is a multilinear function of its columns.*

(b) Use the relation $\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$ to generalize the formula in part (a) to

$$D(\det)(\mathbf{A})\mathbf{H} = \det(\mathbf{A}) \cdot \mathrm{tr}(\mathbf{A}^{-1}\mathbf{H}) \qquad \text{for any} \qquad \mathbf{A} \in \mathrm{GL}(n, \mathbb{F}).$$

(c) Prove that $\mathrm{SL}(n, \mathbb{F})$ is a smooth submanifold of $\mathbb{F}^{n \times n}$, compute its dimension, and show

$$T_{\mathbb{1}} \mathrm{SL}(n, \mathbb{F}) = \left\{ \mathbf{H} \in \mathbb{F}^{n \times n} \;\middle|\; \mathrm{tr}(\mathbf{H}) = 0 \right\}.$$

(d) Consider the set of *non-invertible* n-by-n matrices,

$$M := \left\{ \mathbf{A} \in \mathbb{F}^{n \times n} \;\middle|\; \det(\mathbf{A}) = 0 \right\}.$$

Is 0 a regular value of $\det : \mathbb{F}^{n \times n} \to \mathbb{F}$? Is M a submanifold of $\mathbb{F}^{n \times n}$?
Hint: Clearly M *contains the trivial matrix* $0 \in \mathbb{F}^{n \times n}$. *If* M *is a submanifold, what can you say about the tangent space* $T_0 M \subset \mathbb{F}^{n \times n}$? *In how many different directions can you find smooth paths* $\gamma : (-\epsilon, \epsilon) \to \mathbb{F}^{n \times n}$ *through* 0 *that are contained in* M?

Exercise 4.25. The **special orthogonal** and **special unitary** groups are defined as

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}), \qquad \text{and} \qquad \mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C})$$

respectively. Prove:

(a) $\mathrm{SO}(n)$ is an open (and also closed) subset of $\mathrm{O}(n)$, hence it is a smooth submanifold with the same dimension and $T_{\mathbb{1}} \mathrm{SO}(n) = T_{\mathbb{1}} \mathrm{O}(n)$.

(b) $\mathrm{SU}(n)$ is a smooth submanifold of $\mathrm{U}(n)$ with $\dim \mathrm{SU}(n) = \dim \mathrm{U}(n) - 1$, and

$$T_{\mathbb{1}} \mathrm{SU}(n) = \left\{ \mathbf{H} \in \mathbb{C}^{n \times n} \;\middle|\; \mathbf{H}^\dagger = -\mathbf{H} \text{ and } \operatorname{tr}(\mathbf{H}) = 0 \right\}.$$

Hint: Use Exercise 4.9 to show that the determinant defines a smooth map $\det : \mathrm{U}(n) \to S^1$*, where* S^1 *in this case denotes the unit circle in* $\mathbb{C}$*. Prove that* 1 *is a regular value of this map.*

Finally, we consider an interesting space of matrices that does not form a group, but is nonetheless a manifold.

EXERCISE 4.26. For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and nonnegative integers m, n and $r \leqslant \min\{m, n\}$, let

$$V_r(m, n, \mathbb{F}) := \left\{ \mathbf{A} \in \mathbb{F}^{m \times n} \;\middle|\; \operatorname{rank}(\mathbf{A}) = r \right\}.$$

By the standard formula relating ranks and kernels, $V_r(m, n, \mathbb{F})$ is the set of all m-by-n matrices $\mathbf{A}$ over $\mathbb{F}$ such that $\dim_{\mathbb{F}} \ker \mathbf{A} = n - r$, and the latter condition is also equivalent to $\dim_{\mathbb{F}} \operatorname{coker} \mathbf{A} = m - r$, where the **cokernel** of $\mathbf{A}$ is defined from its image $\operatorname{im}(\mathbf{A}) \subset \mathbb{F}^m$ as the quotient space $\mathbb{F}^m / \operatorname{im}(\mathbf{A})$.

Given any $\mathbf{M}_0 \in V_r(m, n, \mathbb{F})$, one can find splittings $\mathbb{F}^n = V \oplus K$ and $\mathbb{F}^m = W \oplus C$ such that $K = \ker \mathbf{M}_0$ and $W = \operatorname{im} \mathbf{M}_0$. Regarding any other matrix $\mathbf{M} \in \mathbb{F}^{m \times n}$ as a linear map $\mathbb{F}^n \to \mathbb{F}^m$, these splittings of $\mathbb{F}^n$ and $\mathbb{F}^m$ give rise to a block decomposition

$$\mathbf{M} = \begin{pmatrix} \mathbf{A}(\mathbf{M}) & \mathbf{B}(\mathbf{M}) \\ \mathbf{C}(\mathbf{M}) & \mathbf{D}(\mathbf{M}) \end{pmatrix} : V \oplus K \to W \oplus C,$$

thus defining linear (and therefore smooth) maps $\mathbf{A} : \mathbb{F}^{m \times n} \to \operatorname{Hom}(V, W)$, $\mathbf{B} : \mathbb{F}^{m \times n} \to \operatorname{Hom}(K, W)$, $\mathbf{C} : \mathbb{F}^{m \times n} \to \operatorname{Hom}(V, C)$ and $\mathbf{D} : \mathbb{F}^{m \times n} \to \operatorname{Hom}(K, C)$. By construction, the functions $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ all vanish at $\mathbf{M}_0$, while $\mathbf{A}(\mathbf{M}_0) : V \to W$ is invertible. Observe that the invertible maps in $\operatorname{Hom}(V, W)$ form an open subset; this is true for the same reason that $\mathrm{GL}(n, \mathbb{F})$ is an open subset of $\mathbb{F}^{n \times n}$. We can therefore fix an open neighborhood $\mathcal{O} \subset \mathbb{F}^{m \times n}$ of $\mathbf{M}_0$ such that $\mathbf{A}(\mathbf{M}) : V \to W$ is invertible for all $\mathbf{M} \in \mathcal{O}$, and use this to define two smooth maps $\Phi : \mathcal{O} \to \operatorname{Hom}(K, C)$ and $\Psi : \mathcal{O} \to \mathbb{F}^{n \times n}$ by

$$\Phi(\mathbf{M}) := \mathbf{D}(\mathbf{M}) - \mathbf{C}(\mathbf{M})\mathbf{A}(\mathbf{M})^{-1}\mathbf{B}(\mathbf{M}), \quad \text{and} \quad \Psi(\mathbf{M}) := \begin{pmatrix} \mathbb{1} & -\mathbf{A}(\mathbf{M})^{-1}\mathbf{B}(\mathbf{M}) \\ 0 & \mathbb{1} \end{pmatrix},$$

where in the latter expression we are regarding $\Psi(\mathbf{M})$ as a linear map $\mathbb{F}^n \to \mathbb{F}^n$ and writing its block decomposition with respect to the splitting $\mathbb{F}^n = V \oplus K$.

(a) Show that $\Psi(\mathbf{M}) \in \mathbb{F}^{n \times n}$ is invertible for every $\mathbf{M} \in \mathcal{O}$.
(b) Show that for every $\mathbf{M} \in \mathcal{O}$, the kernel of the matrix product $\mathbf{M}\Psi(\mathbf{M}) : \mathbb{F}^n \to \mathbb{F}^m$ is $\{0\} \oplus \ker \Phi(\mathbf{M}) \subset V \oplus K = \mathbb{F}^n$.
(c) Deduce from parts (a) and (b) that $\mathcal{O} \cap V_r(m, n, \mathbb{F}) = \Phi^{-1}(0)$.
Hint: What is the largest dimension that $\ker \mathbf{M}$ *can have for* $\mathbf{M} \in \mathcal{O}$*?*
(d) Show that $\mathbf{M}_0$ is a regular point of Φ, and deduce from this that $V_r(m, n, \mathbb{F}) \subset \mathbb{F}^{m \times n}$ is a smooth submanifold with

$$T_{\mathbf{M}} V_r(m, n, \mathbb{F}) = \left\{ \mathbf{H} \in \mathbb{F}^{m \times n} \;\middle|\; \mathbf{H}(\ker \mathbf{M}) \subset \operatorname{im} \mathbf{M} \right\}$$

for every $\mathbf{M} \in V_r(m, n, \mathbb{F})$, and

$$\dim V_r(m, n, \mathbb{R}) = mn - (m - r)(n - r), \qquad \dim V_r(m, n, \mathbb{C}) = 2 \dim V_r(m, n, \mathbb{R}).$$

(e) A matrix $\mathbf{M} \in \mathbb{F}^{m \times n}$ is said to have **maximal rank** if its rank is $\min\{m, n\}$, which means it is either injective or surjective. Deduce from the result of part (d) that the set of maximal rank matrices is open and dense in $\mathbb{F}^{m \times n}$.

*The result of this exercise produces what is called a **stratification** of $\mathbb{F}^{m\times n}$, meaning that it decomposes $\mathbb{F}^{m\times n}$ into a collection of smooth submanifolds of various dimensions such that every matrix belongs to exactly one of them.*

5. Vector fields

A **vector field** (*Vektorfeld*) X on a smooth manifold M associates to every point $p \in M$ a vector in the corresponding tangent space,

$$X(p) \in T_pM.$$

For example, on $S^2 \subset \mathbb{R}^3$, the tangent space T_pS^2 is the orthogonal complement of the vector $p \in S^2 \subset \mathbb{R}^3$, thus a vector field associates to each such point another vector that is orthogonal to it. We say that a vector field X is **smooth** if the map

$$M \to TM : p \mapsto X(p)$$

is smooth. The set of all smooth vector fields on M forms a vector space, which we will denote by

$$\mathfrak{X}(M) := \left\{ X \in C^\infty(M, TM) \;\middle|\; X(p) \in T_pM \text{ for every } p \in M \right\}.$$

As with real-valued functions, one can define the **support** (*Träger*) of a vector field X as the closure in M of the set $\{p \in M \mid X(p) \neq 0\}$.

5.1. The flow of a vector field. The most important fact about vector fields on manifolds is that they determine dynamical systems. For a smooth path $\gamma : (a, b) \to M$, the derivative

$$\dot\gamma(t) := \frac{d\gamma}{dt}(t) \in T_{\gamma(t)}M$$

can be defined for each $t \in (a, b)$ as a special case of our definition of *partial* derivatives in §3.4. In important special cases such as when M is a submanifold of $\mathbb{R}^n$, $\dot\gamma(t)$ means exactly what you think it should; more generally, it is the equivalence class $[\gamma_t]$ represented by the reparametrized path $\gamma_t(s) := \gamma(t + s)$ that passes through $\gamma(t)$ at $s = 0$. Given $X \in \mathfrak{X}(M)$, a path $\gamma : (a, b) \to M$ is called a **flow line** or **orbit** of X if it satisfies

$$\dot\gamma(t) = X(\gamma(t)).$$

The following fundamental result translates most of the basic existence/uniqueness theory for ordinary differential equations into the language of differential geometry.

THEOREM 5.1. *For any smooth vector field $X \in \mathfrak{X}(M)$ on a manifold M, there exists a unique open subset $\mathcal{O} \subset \mathbb{R} \times M$ and smooth map*

$$\mathcal{O} \to M : (t, p) \mapsto \varphi^t_X(p),$$

*called the **flow** (Fluss) of X, such that for every $p \in M$, the set*

$$\ell_p := \left\{ t \in \mathbb{R} \;\middle|\; (t, p) \in \mathcal{O} \right\} \subset \mathbb{R}$$

is an open interval containing 0, and

$$\gamma_p : \ell_p \to M : t \mapsto \varphi^t_X(p)$$

is the maximal solution to the initial value problem

$$\dot\gamma(t) = X(\gamma(t)), \qquad \gamma(0) = p.$$

Moreover, if X has compact support, then $\mathcal{O} = \mathbb{R} \times M$.

PROOF. For the most part, this result is proved by choosing local coordinates so as to rewrite the initial value problem in $\mathbb{R}^n$ and then applying standard results from the theory of ODEs. We will merely add a few observations in order to see how this works. First, given $p_0 \in M$, choose a smooth chart $(\mathcal{U}, x)$ with $p_0 \in \mathcal{U}$, which gives rise to a smooth chart $(T\mathcal{U}, Tx)$ on TM. The smoothness of X means that $p \mapsto Tx(X(p)) = (x(p), d_p x(X(p)))$ is a smooth function $\mathcal{U} \to \mathbb{R}^{2n}$, thus in particular, so is the function

$$\Phi : \mathcal{U} \to \mathbb{R}^n : p \mapsto d_p x(X(p)).$$

A path $\gamma : (-\epsilon, \epsilon) \to \mathcal{U}$ with $\gamma(0) = p_0$ will now satisfy $\dot{\gamma}(t) = X(\gamma(t))$ if and only if

$$(x \circ \gamma)'(t) = d_{\gamma(t)} x(\dot{\gamma}(t)) = d_{\gamma(t)} x(X(\gamma(t))),$$

meaning that $\alpha := x \circ \gamma : (-\epsilon, \epsilon) \to x(\mathcal{U}) \subset \mathbb{R}^n$ must be a solution to the initial value problem

$$\dot{\alpha}(t) = F(\alpha(t)), \qquad \alpha(0) = x(p_0), \tag{5.1}$$

where we define $F : x(\mathcal{U}) \to \mathbb{R}^n$ by

$$F(q) := d_{x^{-1}(q)} x(X(x^{-1}(q))) = \Phi \circ x^{-1}(q).$$

This last expression shows that F is a smooth function, so in particular it is Lipschitz, and the Picard-Lindelöf theorem therefore applies, telling us that a solution $\alpha : (-\epsilon, \epsilon) \to x(\mathcal{U})$ to (5.1) exists for some $\epsilon > 0$ and is unique. Since F is smooth, this solution also depends smoothly on the initial point $x(p_0)$. Replacing α with $\gamma = x^{-1} \circ \alpha : (-\epsilon, \epsilon) \to \mathcal{U}$, we similarly obtain existence and uniqueness of a solution to $\dot{\gamma}(t) = X(\gamma(t))$ with $\gamma(0) = p_0$, along with smooth dependence on the point p_0. This uniquely defines the flow map $(t, p) \mapsto \varphi_X^t(p)$ for all (t, p) in some neighborhood of $\{0\} \times M \subset \mathbb{R} \times M$.

It remains to establish that the flow map has a unique extension to a maximal domain which is an open subset $\mathcal{O} \subset \mathbb{R} \times M$, and is all of $\mathbb{R} \times M$ if X has compact support. This follows via the same tricks that are used to prove the corresponding statement in $\mathbb{R}^n$, e.g. whenever a flow line $\gamma : [0, T] \to M$ with $\gamma(0) = p_0$ exists, one can find a finite partition $0 = t_0 < t_1 < \ldots < t_{N-1} < t_N = T$ such that the subintervals $[t_{j-1}, t_j]$ are each sufficiently small for $\gamma([t_{j-1}, t_j])$ to lie within the domain of a single chart. One can then make use of the formula

$$\gamma(T) = \varphi_X^T(p_0) = \varphi_X^{t_N - t_{N-1}} \circ \ldots \circ \varphi_X^{t_2 - t_1} \circ \varphi_X^{t_1}(p_0),$$

in which each map in the composition is already known to be smooth and defined on an open neighborhood of the relevant point as long as the increments $t_j - t_{j-1}$ are small enough. This establishes that $\mathcal{O} \subset \mathbb{R} \times M$ is open and $(t, p) \mapsto \varphi_X^t(p)$ is smooth. Finally, if the support $K \subset M$ of X is a compact subset, then clearly every flow line through a point $p_0 \in M \backslash K$ is constant, so that $(t, p_0) \in \mathcal{O}$ for all $t \in \mathbb{R}$. For the same reason, uniqueness of solutions implies that a flow line with initial value at a point $p_0 \in K$ can never escape from K; if it did, then it would become constant outside of K, and must therefore have *always* been a constant path outside of K. We claim now that for every $p_0 \in K$, the maximal solution to $\dot{\gamma}(t) = X(\gamma(t))$ with $\gamma(0) = p_0$ is defined for all $t \in \mathbb{R}$. If not, then suppose $\gamma : (a, b) \to M$ is the maximal solution and either $a > -\infty$ or $b < \infty$; for concreteness we will assume the latter since there is no substantial difference between the two cases. Then (a, b) contains a sequence t_j with $t_j \to b$, and after restricting to a subsequence, the compactness of K implies that we can assume $\gamma(t_j)$ converges to some point $p_1 \in K$. But solutions to the initial value problem starting at points near p_1 also exist and are unique on some sufficiently small interval, so for j large enough, $\gamma(t_j)$ must eventually lie on one of these solutions. The only way to have $\gamma(t_j) \to p_1$ is then if γ eventually matches (up to parametrization) the unique flow line through p_1, in which case it must reach that point at time $t = b$ and can be continued past it; this contradicts the assumption that γ could not be extended beyond the interval (a, b). $\square$

We say that $X \in \mathfrak{X}(M)$ admits a **global flow** if the domain $\mathcal{O} \subset \mathbb{R} \times M$ of the flow map $(t, p) \mapsto \varphi_X^t(p)$ is $\mathbb{R} \times M$. This can sometimes be true even if X does not have compact support, e.g. it is easy to show that every C^0-bounded smooth vector field on $\mathbb{R}^n$ has a global flow. (There are also easy counterexamples if X is allowed to be unbounded, such as $X(x) := x^2$ on $\mathbb{R}$.) In the general case, φ_X^t defines for each $t \in \mathbb{R}$ a smooth map $\mathcal{O}_X^t \to M$ on the open set

$$\mathcal{O}_X^t := \{p \in M \mid (t, p) \in \mathcal{O}\},$$

and in fact, φ_X^t is a diffeomorphism from $\mathcal{O}_X^t$ to $\mathcal{O}_X^{-t}$, with inverse

$$(\varphi_X^t)^{-1} = \varphi_X^{-t}.$$

In particular, if the flow is global, then $\mathcal{O}_X^t = M$ for each $t \in \mathbb{R}$, and φ_X^t is therefore a diffeomorphism from M to itself. It is also possible however to have $\mathcal{O}_X^t = \varnothing$ for $t \neq 0$, though this cannot happen when t is close to 0. Indeed, it follows directly from the definition that

$$\mathcal{O}_X^s \supset \mathcal{O}_X^t \qquad \text{whenever} \quad 0 \leqslant s \leqslant t \quad \text{or} \quad t \leqslant s \leqslant 0,$$

and short-time existence of solutions also implies

$$\mathcal{O}_X^0 = \bigcup_{t>0} \mathcal{O}_X^t = \bigcup_{t<0} \mathcal{O}_X^t = M.$$

The most important properties of the flow are perhaps

$$\varphi_X^0 = \text{Id}, \qquad \text{and} \qquad \varphi_X^{s+t} = \varphi_X^s \circ \varphi_X^t \quad \text{on} \quad \mathcal{O}_X^s \cap \mathcal{O}_X^t \cap \mathcal{O}_X^{s+t} \text{ for every } s, t \in \mathbb{R},$$

which follow from the uniqueness of solutions to the initial value problem. Whenever the flow is global, this means that the map $t \mapsto \varphi_X^t$ defines a group homomorphism from $\mathbb{R}$ to the group $\text{Diff}(M)$ of diffeomorphisms $M \to M$. This is, in practice, the single easiest way to produce a diffeomorphism on a manifold: one need not write it down explicitly, but can instead often write down an appropriate vector field more-or-less explicitly and deduce the existence of a suitable diffeomorphism via its flow. The following exercise is a demonstration of this technique:

Exercise 5.2. A manifold M is called **connected** (*zusammenhängend*)[24] if for every pair of points $p, q \in M$, there exists a continuous path $\gamma : [0, 1] \to M$ from $\gamma(0) = p$ to $\gamma(1) = q$. Show that under this assumption, there exists a diffeomorphism $\varphi : M \to M$ that is the identity map outside of a compact subset and satisfies $\varphi(p) = q$.
Hint: You should first convince yourself that the path $\gamma : [0, 1] \to M$ can be assumed to be a smooth embedding without loss of generality. (This is obvious if γ happens to lie in the domain of a chart $(\mathcal{U}, x)$ such that $x(\mathcal{U}) \subset \mathbb{R}^n$ is convex, and notice that $\gamma([0, 1]) \subset M$ can always be covered by finitely many such charts.) Then choose a vector field that has a flow line containing this path.

Remark 5.3. If the vector field X is not smooth but is of class C^k for some $k \in \mathbb{N}$, then the proof of Theorem 5.1 above can be adapted to produce a flow map $(t, p) \mapsto \varphi_X^t(p)$ that is also of class C^k. As you may recall from your analysis courses, all bets are off if X is continuous but not C^1: in this case local solutions exist but may not be unique, so the flow cannot be defined.

[24] If you know some topology, you may notice that what we are defining here is actually the notion of a **path-connected** space, and connectedness (without mentioning paths) usually means something else. However, every manifold is *locally* path-connected, so a general theorem from point-set topology (see [Wen23, Theorem 7.19]) implies that connectedness and path-connectedness on a manifold are equivalent conditions.

5.2. Pullbacks and pushforwards. A diffeomorphism

$$\psi : M \to N$$

between two manifolds can be viewed as a way of "translating" all geometric data from M into equivalent geometric data on N or vice versa. The exact mechanism for the translation depends on the kind of data we are talking about: for points $p \in M$, the translation in N is simply $\psi(p) \in N$. For a function $f \in C^\infty(M)$, the equivalent data on N is a function

$$\psi_* f \in C^\infty(N)$$

that has the same value at the equivalent point $\psi(p)$ that f has at the original point p, thus

$$\psi_* f \circ \psi = f, \qquad \text{or equivalently} \qquad \psi_* f = f \circ \psi^{-1}.$$

We call $\psi_* f$ the **pushforward** of f via the diffeomorphism ψ. This process is invertible: one can associate to any $f \in C^\infty(N)$ a **pullback**

$$\psi^* f \in C^\infty(M)$$

via ψ, which takes the same value at p that f takes at $\psi(p)$; the definition is thus

$$\psi^* f = f \circ \psi.$$

To do the same trick with tangent vectors, we need to recall that the tangent map of a diffeomorphism $\psi : M \to N$ is also a diffeomorphism $T\psi : TM \to TN$, one which sends T_pM isomorphically to $T_{\psi(p)}N$ for each $p \in M$. This gives the natural way of "translating" tangent vectors between M and N, so for each $X \in TM$ and $Y \in TN$, we denote

$$\psi_* X := T\psi(X) \in TN, \qquad \psi^* Y := T\psi^{-1}(Y) \in TM.$$

The pushforward of a vector field $X \in \mathfrak{X}(M)$ should then be a vector field

$$\psi_* X \in \mathfrak{X}(N)$$

whose value at $\psi(p)$ for each $p \in M$ is the corresponding translation of the tangent vector $X(p)$, namely $\psi_*(X(p))$. This gives

$$(\psi_* X) \circ \psi = T\psi \circ X, \qquad \text{or equivalently} \qquad \psi_* X = T\psi \circ X \circ \psi^{-1}.$$

The pullback of a vector field $Y \in \mathfrak{X}(N)$ is obtained by inverting this procedure, thus

$$\psi^* Y := T\psi^{-1} \circ Y \circ \psi \in \mathfrak{X}(M).$$

PROPOSITION 5.4. *Suppose $\psi : M \to N$ is a diffeomorphism, $X \in \mathfrak{X}(N)$ is a vector field, and $t \in \mathbb{R}$. Then a point $p \in M$ is in the domain of the flow $\varphi^t_{\psi^* X}$ if and only if $\psi(p)$ belongs to the domain of φ^t_X, and $\psi \circ \varphi^t_{\psi^* X} = \varphi^t_X \circ \psi$.*

PROOF. The result follows from the observation that ψ provides a natural bijective correspondence between the flow lines of X on N and flow lines of $\psi^* X$ on M. Indeed, suppose $a < 0 < b$ and $\gamma : (a, b) \to N$ is a flow line of X, satisfying $\dot\gamma(t) = X(\gamma(t))$ and $\gamma(0) = q := \psi(p)$. Then $\alpha := \psi^{-1} \circ \gamma : (a, b) \to M$ satisfies $\alpha(0) = p$ and

$$\dot\alpha(t) = T\psi^{-1}\big(\dot\gamma(t)\big) = T\psi^{-1}\big(X(\gamma(t))\big) = T\psi^{-1} \circ X \circ \psi(\alpha(t)) = (\psi^* X)(\alpha(t)).$$

Conversely, the same computation implies that if α is a flow line of $\psi^* X$, then $\gamma := \psi \circ \alpha$ is a flow line of X. □

EXERCISE 5.5. For two diffeomorphisms $\psi : M \to N$ and $\varphi : N \to Q$, prove the following relations:

(a) $(\varphi \circ \psi)_* f = \varphi_*(\psi_* f) \in C^\infty(Q)$ for $f \in C^\infty(M)$.
(b) $(\varphi \circ \psi)^* g = \psi^*(\varphi^* g) \in C^\infty(M)$ for $g \in C^\infty(Q)$.

(c) $(\varphi \circ \psi)_* X = \varphi_*(\psi_* X) \in \mathfrak{X}(Q)$ for $X \in \mathfrak{X}(M)$.
(d) $(\varphi \circ \psi)^* Y = \psi^*(\varphi^* Y) \in \mathfrak{X}(M)$ for $Y \in \mathfrak{X}(Q)$.

We will see later that when $\psi : M \to N$ is a diffeomorphism, pullbacks and pushforwards can be defined for any meaningful geometric data one might want to consider on M or N. A special case that arises quite often is where $M = N$ and $\psi : M \to M$ is defined by the flow of a vector field; we will see an example of this in the next lecture when we discuss the Lie derivative of a vector field. It will also be important to know that for certain types (but not all types) of data, either the pushforward *or* the pullback (but not both) can be defined via arbitrary smooth maps $\psi : M \to N$, not only for diffeomorphisms. One example of this is already apparent: for $f \in C^\infty(N)$, the pullback

$$\psi^* f := f \circ \psi \in C^\infty(M)$$

makes sense for any smooth map $\psi : M \to N$, so M and N need not be diffeomorphic. One cannot similarly define pushforwards of functions in this context, since ψ^{-1} might not be defined. We will see many more examples of this phenomenon when we discuss tensors and differential forms.

5.3. Derivations. For real-valued functions $f : M \to \mathbb{R}$, there is no natural notion of "partial derivatives" of f, unless M happens to be an open subset of $\mathbb{R}^n$. It is still natural however to talk about the **directional derivative** (*Richtungsableitung*) of f at a point $p \in M$ with respect to a tangent vector $X \in T_pM$, which is computed by evaluating the differential $df : TM \to \mathbb{R}$ of f on X. A closely related notion is the **Lie derivative** (*Lie-Ableitung*) $\mathcal{L}_X f \in C^\infty(M)$ of f with respect to a vector field $X \in \mathfrak{X}(M)$, which is defined by first pulling back the function via the diffeomorphisms φ_X^t for each $t \in \mathbb{R}$, and then differentiating the resulting smooth family of functions with respect to the parameter t:

$$\mathcal{L}_X f := \left.\frac{d}{dt}(\varphi_X^t)^* f\right|_{t=0} \in C^\infty(M).$$

Lie derivatives with respect to a vector field X can similarly be defined on any geometric objects for which the notion of pulling back via a diffeomorphism makes sense, e.g. we will consider Lie derivatives of a vector field in the next lecture. The Lie derivative of a real-valued function f turns out to be the same thing as computing the directional derivative of f with respect to $X(p) \in T_pM$ at each point $p \in M$: indeed, since $(\varphi_X^t)^* f = f \circ \varphi_X^t$, we have

$$(\mathcal{L}_X f)(p) = \left.\frac{d}{dt} f\left(\varphi_X^t(p)\right)\right|_{t=0} = df\left(\left.\frac{d}{dt}\varphi_X^t(p)\right|_{t=0}\right) = df(X(p)),$$

or in more succinct notation,

$$\mathcal{L}_X f \equiv df(X).$$

Note that this discussion does not require the vector field X to have a global flow: strictly speaking, φ_X^t may not be a globally-defined diffeomorphism for all $t \in \mathbb{R}$, but on any given compact neighborhood of a point, φ_X^t can always be defined for $t \in \mathbb{R}$ sufficiently close to 0, which is good enough for all of the definitions and formulas above to make sense.

The differential operator $\mathcal{L}_X$ associated to any $X \in \mathfrak{X}(M)$ defines a map

$$\mathcal{L}_X : C^\infty(M) \to C^\infty(M) : f \mapsto \mathcal{L}_X f,$$

and one can check using the usual rules of differentiation that this map is linear:

$$\mathcal{L}_X(f + g) = \mathcal{L}_X f + \mathcal{L}_X g, \qquad \mathcal{L}_X(cf) = c\mathcal{L}_X f, \qquad \text{for all } f, g \in C^\infty(M),\ c \in \mathbb{R}.$$

Moreover, the product rule for differentiation translates into the following so-called **Leibniz rule**:

$$\mathcal{L}_X(fg) = (\mathcal{L}_X f)g + f\mathcal{L}_X g.$$

This formula motivates a short digression on algebras and Lie algebras.

DEFINITION 5.6. An **algebra** is a vector space $\mathcal{A}$ that is endowed with the additional structure of a bilinear multiplication operation

$$\mathcal{A} \times \mathcal{A} \to \mathcal{A} : (x, y) \mapsto xy$$

that is also associative, i.e. $(xy)z = x(yz)$ for all $x, y, z \in \mathcal{A}$.[25] A **derivation** on $\mathcal{A}$ is a linear map $L : \mathcal{A} \to \mathcal{A}$ that satisfies the Leibniz rule

$$L(xy) = (Lx)y + x(Ly) \qquad \text{for all } x, y \in \mathcal{A}.$$

An algebra endowed with a derivation is called a **differential algebra** (*Differentialalgebra*).

DEFINITION 5.7. A **Lie algebra** (*Lie-Algebra*) is a vector space V that is endowed with the additional structure of a bilinear operation

$$[\cdot, \cdot] : V \times V \to V,$$

its so-called **Lie bracket** (*Lie-Klammer*), which satisfies:

- **antisymmetry**: $[u, v] = -[v, u]$ for all $u, v \in V$;
- the **Jacobi identity**: $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ for all $u, v, w \in V$.

EXERCISE 5.8. Show that on any algebra $\mathcal{A}$, the space $\mathcal{D}$ of all derivations on $\mathcal{A}$ can be made into a Lie algebra by defining the bracket

$$[L_1, L_2] := L_1 \circ L_2 - L_2 \circ L_1.$$

In this course, the most important example of an algebra is the space of smooth real-valued functions $C^\infty(M)$ on a manifold M, in which multiplication is defined pointwise by $(fg)(p) := f(p)g(p)$. The previous remarks show that for any smooth vector field $X \in \mathfrak{X}(M)$, the associated Lie derivative operator $\mathcal{L}_X$ defines a derivation on $C^\infty(M)$. A somewhat less obvious class of examples comes from the observation in Exercise 5.8 that the **commutator bracket** of any two derivations is also a derivation, so in particular, any pair of vector fields $X, Y \in \mathfrak{X}(M)$ gives rise to a derivation on $C^\infty(M)$ defined by

$$[\mathcal{L}_X, \mathcal{L}_Y]f = \mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f.$$

One says that the vector fields X and Y **commute** (*kommutieren*) whenever this bracket vanishes. This will turn out to be an important condition, but its meaning will take some effort to unpack. We first need to make the surprising and useful observation that the examples we have seen so far of derivations on $C^\infty(M)$ are the *only* examples that exist:

THEOREM 5.9. *Every derivation $L : C^\infty(M) \to C^\infty(M)$ is of the form $L = \mathcal{L}_X$ for some (unique) smooth vector field $X \in \mathfrak{X}(M)$.*

PROOF. The uniqueness of X is clear, since different vector fields define different derivations. The proof of existence follows from a series of claims.

***Claim 1**: If $f : M \to \mathbb{R}$ is a constant function, then $Lf = 0$ for every derivation L on $C^\infty(M)$.*

Indeed, if f is constant, then multiplication of an arbitrary function $g \in C^\infty(M)$ by f is the same as scalar multiplication, so linearity implies $L(fg) = f\,Lg$, and combining this with the Leibniz rule gives $(Lf)g = 0$. Plugging in the function $g \equiv 1$, we conclude $Lf \equiv 0$.

[25] If you're into algebra, you may notice that the definition of an algebra is quite similar to that of a ring. The difference is that while a ring is also an abelian group with respect to its "+" operation and has a distributive product operation, it does not generally come with any notion of scalar multiplication and is thus not a vector space. One can however define the notion of an algebra more generally, so that it is a module over a commutative ring R instead of a vector space. The case where R is a field then agrees with the definition we've given, but one can also speak of an algebra over $\mathbb{Z}$, which is the same thing as a ring since modules over $\mathbb{Z}$ are the same thing as abelian groups.

Claim 2: *The stated result is true in the special case where M is a convex open subset of Euclidean space, $\Omega \subset \mathbb{R}^n$.*

This is the heart of the proof, and it depends on an important fact in first-year analysis that follows from the fundamental theorem of calculus. Assume $\Omega \subset \mathbb{R}^n$ is open and convex, and fix a point $x_0 = (x_0^1, \dots, x_0^n) \in \Omega$. For any other point $x = (x^1, \dots, x^n) \in \Omega$, the convexity of Ω implies that it contains the line segment between x_0 and x, so using the fundamental theorem of calculus and the chain rule, we find that any smooth function $f : \Omega \to \mathbb{R}$ satisfies

$$\begin{aligned} f(x) &= f(x_0) + \int_0^1 \frac{d}{d\tau} f(x_0 + \tau(x - x_0))\, d\tau = f(x_0) + \int_0^1 Df(x_0 + \tau(x - x_0))(x - x_0)\, d\tau \\ &= f(x_0) + \sum_{j=1}^n \left(\int_0^1 \partial_j f(x_0 + \tau(x - x_0))\, d\tau \right) (x^j - x_0^j) =: f(x_0) + \sum_{j=1}^n h_j(x)(x^j - x_0^j), \end{aligned} \tag{5.2}$$

where we've defined smooth functions $h_j : \Omega \to \mathbb{R}$ by $h_j(x) := \int_0^1 \partial_j f(x_0 + \tau(x - x_0))\, d\tau$. To make use of this formula, we can regard each of the coordinates $x^1, \dots, x^n$ as smooth real-valued functions on Ω and associate to these the smooth functions

$$X^j := L(x^j) \in C^\infty(\Omega), \qquad j = 1, \dots, n.$$

Linearity and the Leibniz rule, together with Claim 1, now produce from (5.2) the formula $Lf(x) = \sum_{j=1}^n \left[Lh_j(x) \cdot (x^j - x_0^j) + h_j(x) X^j(x) \right]$, so in particular,

$$Lf(x_0) = \sum_{j=1}^n h_j(x_0) X^j(x_0) = \sum_{j=1}^n X^j(x_0) \partial_j f(x_0).$$

The definition of the functions $X^j \in C^\infty(\Omega)$ did not depend on the choice of point $x_0 \in \Omega$, thus this formula is valid for every such point, giving an equality of functions

$$Lf = \sum_{j=1}^n X^j \partial_j f = \mathcal{L}_X f \qquad \text{on } \Omega,$$

where we define the smooth vector field $X \in \mathfrak{X}(\Omega)$ by $X(x) = (X^1(x), \dots, X^n(x)) \in \mathbb{R}^n = T_x\Omega$.

Claim 3: *If the theorem holds for a particular manifold M, then it also holds for every manifold that is diffeomorphic to M.*

Assume $\psi : N \to M$ is a diffeomorphism between two manifolds, and the theorem is already known to hold for M. Any derivation L on $C^\infty(N)$ then determines a "pushforward" derivation $\psi_* L$ on $C^\infty(M)$ via the formula

$$(\psi_* L) f := L(f \circ \psi) \circ \psi^{-1}. \tag{5.3}$$

By assumption, the latter is $\mathcal{L}_X$ for some vector field $X \in \mathfrak{X}(M)$, and it is reasonable to guess that L will therefore correspond to the pullback vector field $\psi^* X \in \mathfrak{X}(N)$ as defined in §5.2. Let's check this: $\psi^* X$ is defined by

$$\psi^* X(p) = T\psi^{-1}(X(\psi(p))).$$

For $g \in C^\infty(N)$ and $p \in N$, we define $f := g \circ \psi^{-1} \in C^\infty(M)$ and use (5.3) to write

$$\begin{aligned} (Lg)(p) &= L(f \circ \psi)(p) = [(\psi_* L) f](\psi(p)) = (\mathcal{L}_X f)(\psi(p)) = df(X(\psi(p))) \\ &= d(g \circ \psi^{-1})(X(\psi(p))) = dg \circ T\psi^{-1}(X(\psi(p))) = dg(\psi^* X(p)) = \mathcal{L}_{\psi^* X} g(p), \end{aligned}$$

so the guess is correct!

For the remaining claims, assume M is a fixed manifold and $L : C^\infty(M) \to C^\infty(M)$ is a derivation.

Claim 4: *If $f \in C^\infty(M)$ vanishes on a neighborhood of some point $p \in M$, then $Lf(p) = 0$.*

To see this, suppose $\mathcal{U} \subset M$ is a neighborhood of p on which $f \in C^\infty(M)$ vanishes, and choose any $g \in C^\infty(M)$ so that $g(p) = 1$ but g has compact support in $\mathcal{U}$.[26] Then $fg = 0$, thus $0 = (Lf)g + f(Lg)$, and evaluating the right hand side at p gives $0 = Lf(p) \cdot g(p) = Lf(p)$.

In light of linearity, a corollary of Claim 4 is that for any $f \in C^\infty(M)$, the value of $Lf(p)$ at any given point $p \in M$ depends only on the values of f on an arbitrarily small neighborhood of p.

Claim 5: *For any open subset $\mathcal{U} \subset M$, L determines a unique derivation $L_\mathcal{U} : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ such that for every $f \in C^\infty(M)$, $L_\mathcal{U}(f|_\mathcal{U}) = (Lf)|_\mathcal{U}$.*

This follows from the observation at the end of Claim 4 that $Lf(p)$ depends on f only in a neighborhood of p. Indeed, for any $f \in C^\infty(\mathcal{U})$, there is a unique function $L_\mathcal{U} f \in C^\infty(\mathcal{U})$ characterized by the property that for each $p \in \mathcal{U}$ and $f_p \in C^\infty(M)$ with $f_p \equiv f$ near p, $L_p f \equiv L_\mathcal{U} f$ near p. It is straightforward to verify that $L_\mathcal{U}$ defined in this way is a derivation.

Conclusion: Choose an open cover $M = \bigcup_{\alpha \in I} \mathcal{U}_\alpha$ such that for every α, there is a chart $(\mathcal{U}_\alpha, x_\alpha)$ whose image $x(\mathcal{U}_\alpha) \subset \mathbb{R}^n$ is convex. Claims 2 and 3 imply that the theorem holds for each of the open subsets $\mathcal{U}_\alpha \subset M$, thus for the derivation L_α determined on $C^\infty(\mathcal{U}_\alpha)$ by Claim 5, we have $L_\alpha = \mathcal{L}_{X_\alpha}$ for some vector field $X_\alpha \in \mathfrak{X}(\mathcal{U}_\alpha)$. We claim that for every pair $\alpha, \beta \in I$, X_α and X_β match on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. Indeed, if $X_\alpha(p) \neq X_\beta(p)$ for some point p, then we can find a function $f \in C^\infty(M)$ with compact support in $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ such that $\mathcal{L}_{X_\alpha} f(p) \neq \mathcal{L}_{X_\beta} f(p)$, which is a contradiction since $L_\alpha(f|_{\mathcal{U}_\alpha})$ and $L_\beta(f|_{\mathcal{U}_\beta})$ should both have the same restriction as Lf on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. The claim now implies that the vector fields X_α can be patched together to form a smooth vector field $X \in \mathfrak{X}(M)$, and in light of Claim 4, the relation $Lf = \mathcal{L}_X f$ now follows on each $\mathcal{U}_\alpha$ from $L(f|_{\mathcal{U}_\alpha}) = \mathcal{L}_{X_\alpha}(f|_{\mathcal{U}_\alpha})$. ∎

Remark 5.10. In light of Theorem 5.9, it is common in differential geometry to blur the distinction between smooth vector fields on M and derivations on $C^\infty(M)$, and many books even use exactly the same notation for both, thus writing

$$Xf := \mathcal{L}_X f \in C^\infty(M)$$

so as to view the vector field $X \in \mathfrak{X}(M)$ as a differential operator acting on the function $f \in C^\infty(M)$. I personally prefer not to do this, and will thus continue writing $\mathcal{L}_X$ to distinguish the derivation defined by a vector field $X \in \mathfrak{X}(M)$ from the vector field itself; the sole exception to this will be the coordinate vector fields discussed in the next subsection. Many authors would probably call this practice overly pedantic, and I cannot say with confidence that they are wrong.

Exercise 5.11. For a diffeomorphism $\psi : M \to N$, vector field $X \in \mathfrak{X}(M)$ and function $f \in C^\infty(M)$, prove $\mathcal{L}_{\psi_* X}(\psi_* f) = \psi_*(\mathcal{L}_X f) \in C^\infty(N)$.

6. The Lie algebra of vector fields

We saw in the last lecture that there is a natural equivalence between the space of smooth vector fields $\mathfrak{X}(M)$ on a smooth manifold M and the space of all derivations $L : C^\infty(M) \to C^\infty(M)$ on the algebra of smooth functions. It was also observed in Exercise 5.8 that the latter has a natural Lie algebra structure defined via the commutator bracket

$$[L_1, L_2] := L_1 L_2 - L_2 L_1,$$

which is antisymmetric and satisfies the Jacobi identity (see Definition 5.7). Lie algebras are a large topic, and if you have not seen them at all before, then I would not expect you to have any intuition as to why a bilinear bracket satisfying antisymmetry and the Jacobi identity might be an

[26]Such a function can be constructed in local coordinates our of functions of the form $\mathbb{R}^n \to [0,1] : x \mapsto \beta(|x|^2)$, where $\beta : \mathbb{R} \to [0,1]$ is a smooth function with $\beta(t) = 0$ for all $t \geqslant \epsilon > 0$ and $\beta(0) = 1$. The construction of β is an easy exercise once you've seen examples like $h(t) := e^{-1/t^2}$, a smooth function on $(0,\infty)$ admitting a smooth extension to $\mathbb{R}$ that vanishes on $(-\infty, 0]$.

interesting or useful object to study. We will see a first example of the answer to that question in this lecture: the Lie algebra structure on the space of vector fields characterizes the commutativity (or lack thereof) of their respective flows. This will be easily the deepest result we have proved so far in this course, and it will serve as a foundation for several later results involving curvature and integrability.

6.1. Coordinate vector fields. Given a smooth chart $(\mathcal{U}, x)$ on a manifold M, the coordinate functions $x^1, \ldots, x^n : \mathcal{U} \to \mathbb{R}$ define a natural family of derivations on $C^\infty(\mathcal{U})$, namely the n partial derivative operators

$$\partial_j := \frac{\partial}{\partial x^j} : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U}), \qquad j = 1, \ldots, n,$$

which are defined by writing any function $f \in C^\infty(\mathcal{U})$ in its local coordinate representation $(x^1, \ldots, x^n) \mapsto f(x^1, \ldots, x^n)$ and differentiating the resulting function of n variables as one would in first-year analysis. The more precise way to say this is that for each $f \in C^\infty(\mathcal{U})$ and $p \in \mathcal{U}$, the function $\partial_j f \in C^\infty(\mathcal{U})$ is given by

$$(\partial_j f)(p) := \partial_j (f \circ x^{-1})(x(p)),$$

where the right-hand side is a perfectly ordinary partial derivative of a real-valued function of n real variables. The fact that the operators $\partial_1, \ldots, \partial_n$ define derivations on $C^\infty(\mathcal{U})$ follows immediately from the usual product rule. The corresponding vector fields in $\mathfrak{X}(\mathcal{U})$ are also easy to identify: they come from the standard basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$ as transferred over to $\mathcal{U}$ by the chart, i.e. the derivation ∂_j corresponds to the vector field

$$v_j(p) := (d_p x)^{-1}(e_j), \qquad p \in \mathcal{U}.$$

Since this notation is bit clumsy, it has become conventional in differential geometry to use the notation

$$\partial_1, \ldots, \partial_n \quad \text{or equivalently} \quad \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n} \in \mathfrak{X}(\mathcal{U})$$

not just for the derivations but also for the corresponding vector fields on $\mathcal{U}$, and I will follow that convention in these notes, in spite of what I said in Remark 5.10 above. We call these the **coordinate vector fields** determined on $\mathcal{U}$ by the chart $(\mathcal{U}, x)$. Two issues are very important to understand:

(1) The vector fields $\frac{\partial}{\partial x^j}$ are *only* defined on $\mathcal{U} \subset M$; it does not make sense to write down formulas involving ∂_j everywhere on M unless $(\mathcal{U}, x)$ happens to be a *global* chart, meaning $\mathcal{U} = M$.
(2) For each individual $j \in \{1, \ldots, n\}$, the vector field $\frac{\partial}{\partial x^j}$ depends not only on the coordinate function $x^j : \mathcal{U} \to \mathbb{R}$ but on all n of the coordinates $x^1, \ldots, x^n$. Indeed, the vector $\frac{\partial}{\partial x^j}$ points in the unique direction where x^j increases but all the other coordinates are constant. The issue is easy to see in simple examples, e.g. using the standard polar coordinates (r, θ) and Cartesian coordinates (x, y) on suitable regions in $\mathbb{R}^2$, one can define both (r, θ) and (r, y) as smooth charts on the open right half-plane $\{x > 0\} \subset \mathbb{R}^2$. But the partial derivative operator $\frac{\partial}{\partial r}$ has different meanings in these two coordinate systems, because differentiating in a direction where r increases but θ is constant does not typically give the same result as differentiating in a direction where r increases but y is constant.

6.2. Components and the summation convention. Since the coordinate vector fields $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \in \mathfrak{X}(\mathcal{U})$ determined by a chart $(\mathcal{U}, x = (x^1, \dots, x^n))$ on M form a basis of T_pM at each point $p \in \mathcal{U}$, any $X \in \mathfrak{X}(M)$ restricted to $\mathcal{U} \subset M$ can be written uniquely in the form

$$X = \sum_{i=1}^n X^i \partial_i = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i} \tag{6.1}$$

for uniquely defined smooth functions $X^1, \dots, X^n \in C^\infty(\mathcal{U})$, called the **components** of X with respect to the chart $(\mathcal{U}, x)$. This observation will be useful for computations, but it becomes more so if we can make the notation a bit less cumbersome. Einstein introduced a nice trick for this, which is known as the **Einstein summation convention**: the trick is to omit the summation symbol, but assume that whenever a matching pair of "upper" and "lower" indices appears, a summation of that index over all coordinates (in this case from 1 to n) is implied. Using this convention, (6.1) becomes

$$X = X^i \partial_i = X^i \frac{\partial}{\partial x^i},$$

where the convention is also to interpret the upper index in $\frac{\partial}{\partial x^i}$ as a lower index because it appears in the denominator. (I advise you not to search for any deeper meaning behind this—just take it as a definition for now, and you will see presently why it is useful.) The simplicity of this expression in comparison with (6.1) is perhaps not so dramatic, but the Einstein convention becomes especially useful in situations where multiple indices need to be summed over at the same time, which will happen a lot once we start talking about tensors next week.

Let us derive a coordinate transformation formula: suppose $(\widetilde{\mathcal{U}}, \widetilde{x})$ is a second chart with $\mathcal{U} \cap \widetilde{\mathcal{U}} \neq \varnothing$, and the components of X in these alternative coordinates over $\widetilde{\mathcal{U}}$ are denoted by $\widetilde{X}^i$, so $X = \widetilde{X}^i \frac{\partial}{\partial \widetilde{x}^i}$ on $\widetilde{\mathcal{U}}$. How do the components X^i and $\widetilde{X}^i$ relate to each other on the region $\mathcal{U} \cap \widetilde{\mathcal{U}}$ where their domains overlap?

To answer this, we start with the observation that for any $f \in C^\infty(\mathcal{U} \cap \widetilde{\mathcal{U}})$, the chain rule relates the partial derivatives of f with respect to the two different coordinate systems by

$$\frac{\partial f}{\partial x^i} = \frac{\partial f}{\partial \widetilde{x}^j} \frac{\partial \widetilde{x}^j}{\partial x^i}, \tag{6.2}$$

where the Einstein convention gives an implied summation $\sum_{j=1}^n$ on the right hand side. This formula is hopefully familiar to you from analysis, at least when applied to functions on open subsets of $\mathbb{R}^n$; in the present setting, the partial derivatives on both sides are interpreted as derivations applied to smooth functions on $\mathcal{U} \cap \widetilde{\mathcal{U}} \subset M$, but these have been defined in terms of ordinary partial derivatives of functions on $\mathbb{R}^n$. In that context, the left hand side is the ith component of the *gradient* ∇f of f in coordinates $(x^1, \dots, x^n)$, interpreted as a row vector, while the right hand side is the ith component of the product of the row vector $\widetilde{\nabla} f$ (the gradient of f is coordinates $(\widetilde{x}^1, \dots, \widetilde{x}^n)$ with the Jacobian matrix $\frac{\partial \widetilde{x}}{\partial x}$ of the transition map $(x^1, \dots, x^n) \mapsto (\widetilde{x}^1(x^1, \dots, x^n), \dots, \widetilde{x}^n(x^1, \dots, x^n))$. Equation (6.2) is thus equivalent to the relation

$$D(f \circ x^{-1})(x(p)) = D(f \circ \widetilde{x}^{-1})(\widetilde{x}(p)) \circ D(\widetilde{x} \circ x^{-1})(x(p)),$$

which follows directly from the chain rule. Now, the function f was not actually important in this discussion at all: what we are really interested in is a formula relating derivations, namely

$$\frac{\partial}{\partial x^i} = \frac{\partial \widetilde{x}^j}{\partial x^i} \frac{\partial}{\partial \widetilde{x}^j}, \tag{6.3}$$

which can now equally well be interpreted as a formula for the coordinate vector field $\frac{\partial}{\partial x^i}$ as a linear combination of the other set of coordinate vector fields $\frac{\partial}{\partial \widetilde{x}^j}$ where they overlap. This implies

$$X = X^i \frac{\partial}{\partial x^i} = X^i \frac{\partial \widetilde{x}^j}{\partial x^i} \frac{\partial}{\partial \widetilde{x}^j} = \widetilde{X}^j \frac{\partial}{\partial \widetilde{x}^j},$$

from which we derive (after interchanging the indices i and j just for good measure) the transformation formula

$$\widetilde{X}^i = \frac{\partial \widetilde{x}^i}{\partial x^j} X^j. \tag{6.4}$$

You may agree that if we'd had to write summation symbols in all of these expressions, we would be slightly more tired now. Notice that this formula has an easy interpretation in terms of matrix-vector multiplication: if we package the components together into $\mathbb{R}^n$-valued functions $\xi := (X^1, \dots, X^n) : \mathcal{U} \to \mathbb{R}^n$ and $\widetilde{\xi} := (\widetilde{X}^1, \dots, \widetilde{X}^n) : \widetilde{\mathcal{U}} \to \mathbb{R}^n$, then (6.4) relates these two functions to each other via multiplication with the Jacobian matrix $\frac{\partial \widetilde{x}}{\partial x}$:

$$\widetilde{\xi} = \frac{\partial \widetilde{x}}{\partial x} \xi.$$

The Einstein convention has nothing intrinsically to do with differential geometry—it is actually just linear algebra. Once you get used to it, you may begin to wish you had always been doing linear algebra this way.

We will use the Einstein convention consistently throughout the rest of this course, and only include explicitly written summation symbols in situations where their omission might cause confusion.

Remark 6.1. Using the summation convention requires being very careful and consistent about the distinction between upper and lower indices: coordinates and components of vector fields are *always* written with upper indices, while partial derivative operators (and their associated coordinate vector fields) always carry lower indices. Forgetting these conventions can cause grave confusion and should be avoided at all costs. Unfortunately, not all differential geometry books written by mathematicians are completely consistent about this, though books by physicists are—Einstein was one of them, after all, so his mathematical innovations are taken as gospel.

6.3. The Lie bracket. The **Lie bracket** (*Lie-Klammer*) of two vector fields $X, Y \in \mathfrak{X}(M)$ on a manifold M is defined to be the unique vector field

$$[X, Y] \in \mathfrak{X}(M) \qquad \text{such that} \qquad \mathcal{L}_{[X,Y]} = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X.$$

This definition makes sense as a consequence of Exercise 5.8 and Theorem 5.9. In particular, we say that X and Y **commute** if $[X, Y] \equiv 0$.

Exercise 6.2. Suppose $(\mathcal{U}, x)$ is a chart on M and we express two vector fields $X, Y \in \mathfrak{X}(M)$ over $\mathcal{U}$ in this chart as $X = X^i \partial_i$ and $Y = Y^i \partial_i$.

(a) Show that the components $[X, Y]^i$ of $[X, Y]$ with respect to the same chart are given by

$$[X, Y]^i = X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}. \tag{6.5}$$

(b) Use the coordinate transformation formulas (6.3) and (6.4) to give a direct computational proof (without using the result of part (a)) that the vector field defined on $\mathcal{U}$ via the right hand side of (6.5) depends only on $X, Y \in \mathfrak{X}(\mathcal{U})$ and not on the choice of chart $(\mathcal{U}, x)$. In other words, show that for any other chart $(\widetilde{\mathcal{U}}, \widetilde{x})$,

$$\left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right) \frac{\partial}{\partial x^i} = \left(\widetilde{X}^j \frac{\partial \widetilde{Y}^i}{\partial \widetilde{x}^j} - \widetilde{Y}^j \frac{\partial \widetilde{X}^i}{\partial \widetilde{x}^j} \right) \frac{\partial}{\partial \widetilde{x}^i} \quad \text{on} \quad \mathcal{U} \cap \widetilde{\mathcal{U}}.$$

Hint: The matrices with entries $\frac{\partial \tilde{x}^i}{\partial x^j}$ *and* $\frac{\partial x^i}{\partial \tilde{x}^j}$ *are Jacobi matrices for transformations that are inverse to each other, thus they satisfy*

$$\frac{\partial \tilde{x}^i}{\partial x^j}\frac{\partial x^j}{\partial \tilde{x}^k} = \delta^i_k := \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

REMARK 6.3. Physicists like being able to do explicit computations, so they tend to emphasize coordinate-based formulas in this subject much more than mathematicians do. For example, some physics books take the formula (6.5) as a *definition* of the Lie bracket $[X, Y]$, without first talking about commutators of derivations. The price for doing this is that one must prove that switching to a different local coordinate system would not change the definition, i.e. one must do Exercise 6.2(b). The exercise is tedious, but I recommend doing it exactly once in your life, as it may give you some useful insight into the way that physicists do mathematics, and in any case, it is never bad to get better at explicit computations. As a cautionary tale, I also recommend convincing yourself that the simpler formula

$$X^j \frac{\partial Y^i}{\partial x^j}\frac{\partial}{\partial x^i} = \tilde{X}^j \frac{\partial \tilde{Y}^i}{\partial \tilde{x}^j}\frac{\partial}{\partial \tilde{x}^i} \quad \text{on} \quad \mathcal{U} \cap \tilde{\mathcal{U}}$$

is *false* in general, thus one cannot define a vector field $Z = Z^i \partial_i$ by $Z^i := X^j \partial_j Y^i$ and expect the definition to be independent of the choice of coordinates.

EXERCISE 6.4. For $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, give two proofs of the formulas

$$[fX, Y] = f[X, Y] - (\mathcal{L}_Y f)X, \qquad [X, fY] = f[X, Y] + (\mathcal{L}_X f)Y,$$

using different methods:

(a) Directly from the definition of the Lie bracket via Theorem 5.9;
(b) Using the coordinate formula (6.5).

EXERCISE 6.5. For a diffeomorphism $\psi : M \to N$ and two vector fields $X, Y \in \mathfrak{X}(M)$, prove $\psi_*[X, Y] = [\psi_* X, \psi_* Y] \in \mathfrak{X}(N)$.

EXAMPLE 6.6. The coordinate vector fields $\partial_1, \ldots, \partial_n$ defined from any chart on an open subset all commute with each other. One can deduce this either from the fact that $\partial_i \partial_j f = \partial_j \partial_i f$ for all smooth functions f,[27] or as a trivial application of the formula in Exercise 6.2.

My goal for the rest of this lecture is to explain not just what the Lie bracket of two vector fields *is*, but what it *means*. The discussion starts with the following observation related to Example 6.6 above. Consider the manifold $M = \mathbb{R}^n$ with the standard Cartesian coordinates $x^1, \ldots, x^n$ regarded as a global chart on M; this chart is actually just the identity map $\mathbb{R}^n \to \mathbb{R}^n$. The resulting coordinate vector fields $\partial_1, \ldots, \partial_n$ produce the standard basis of the tangent space $T_p\mathbb{R}^n = \mathbb{R}^n$ at every point $p \in \mathbb{R}^n$. It is easy to write down the flow of ∂_j for each $j = 1, \ldots, n$: it is

$$\varphi^t_{\partial_j}(x^1, \ldots, x^n) = (x^1, \ldots, x^{j-1}, x^j + t, x^{j+1}, \ldots, x^n).$$

We see from this that for any two $i, j \in \{1, \ldots, n\}$ and $s, t \in \mathbb{R}$, the corresponding flows commute:

$$\varphi^s_{\partial_i} \circ \varphi^t_{\partial_j} = \varphi^t_{\partial_j} \circ \varphi^s_{\partial_i}.$$

This is a generalization of the basic observation that if you start from some point (x, y) in the plane $\mathbb{R}^2$, move a distance s to the right and then a distance t upward, you'll end up at the same point as if you had made those two moves in the reverse order, namely $(x+s, y+t)$. In other words,

[27]And since this is not an analysis course, there is no need to worry about the fact that $\partial_i \partial_j f = \partial_j \partial_i f$ does not generally hold for functions whose second-order derivatives exist but are discontinuous. With very few exceptions, all functions that we choose to worry about in the remainder of this course will be of class C^∞.

the two paths, each consisting of two straight line segments, combine to form a closed rectangle. This observation is not as trivial as it may seem: in particular, it becomes false in general if you replace ∂_i and ∂_j by different vector fields, e.g. in the example of $\mathbb{R}^2$, one could replace the "horizontal" coordinate vector field ∂_1 with one that still points in the x-direction but flows at different speeds along the lower and upper segments of the rectangle, in which case the rectangle fails to close up. There is no reason in general why the flows of two vector fields should always commute. They do commute in the case of coordinate vector fields on $\mathbb{R}^n$, and it follows easily that flows of coordinate vector fields determined by a chart $(\mathcal{U}, x)$ on a manifold M will generally commute as long as one keeps s and t close enough to 0 so that the flow lines do not escape from $\mathcal{U}$. But pairs of coordinate vector fields are special, and one symptom of this is the fact that their Lie brackets vanish. We will show in §6.5 that this is a general phenomenon: in particular, for any two vector fields $X, Y \in \mathfrak{X}(M)$ whose flows exist globally, one has $\varphi_X^s \circ \varphi_Y^t = \varphi_Y^t \circ \varphi_X^s$ for all $s, t \in \mathbb{R}$ if and only if $[X, Y] \equiv 0$.

6.4. The Lie derivative of a vector field. Before we can prove a result on commuting flows, we need a short digression to address the following question: What might it mean to differentiate a vector field $Y \in \mathfrak{X}(M)$ at a point $p \in M$ in the direction $X \in T_pM$? A naive attempt to define this would proceed as follows: choose any smooth path $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ and $\dot{\gamma}(0) = X$, and set

$$\mathcal{L}_X Y(p) \stackrel{?}{:=} \frac{d}{dt} Y(\gamma(t))\Big|_{t=0} = \lim_{t \to 0} \frac{Y(\gamma(t)) - Y(p)}{t} \;?$$

If Y were a real-valued function instead of a vector field, then we would be on solid ground with this definition, but for a vector field the right hand side does not make sense: outside of the uninteresting special case where γ is a constant path, $Y(\gamma(t)) \in T_{\gamma(t)}M$ and $Y(p) \in T_pM$ generally belong to different vector spaces, so there is no well-defined way of subtracting one from the other.

A solution to this conundrum arises if one allows X to be a vector *field* on M, rather than just a single tangent vector. In this case, the flow of X gives a natural choice of the path

$$\gamma(t) = \varphi_X^t(p),$$

which is defined for t in a sufficiently small interval $(-\epsilon, \epsilon)$ even if the flow does not globally exist. More importantly, the tangent map of the flow gives rise to natural isomorphisms,

$$T_p\varphi_X^t : T_pM \to T_{\varphi_X^t(p)}M = T_{\gamma(t)}M$$

for t close to 0, which gives us a way of identifying with each other the distinct tangent spaces in which $Y(p)$ and $Y(\gamma(t))$ live. Since the inverse of $T\varphi_X^t$ is $T\varphi_X^{-t}$, it now makes sense to define the **Lie derivative** (*Lie-Ableitung*) of $Y \in \mathfrak{X}(M)$ with respect to $X \in \mathfrak{X}(M)$ as the vector field

$$\mathcal{L}_X Y \in \mathfrak{X}(M), \qquad \mathcal{L}_X Y(p) := \frac{d}{dt} T\varphi_X^{-t}\big(Y(\varphi_X^t(p))\big)\Big|_{t=0} = \lim_{t \to 0} \frac{T\varphi_X^{-t}\big(Y(\varphi_X^t(p))\big) - Y(p)}{t}.$$

Recalling the definition of the *pullback* of a vector field in §5.2, we can abbreviate this formula as

$$\mathcal{L}_X Y = \frac{d}{dt} (\varphi_X^t)^* Y\Big|_{t=0}.$$

It turns out that $\mathcal{L}_X Y$ is just a new perspective on the Lie bracket:

Proposition 6.7. *For any* $X, Y \in \mathfrak{X}(M)$, $\mathcal{L}_X Y = [X, Y]$.

Proof. We need to show that for every $f \in C^\infty(M)$,

$$\mathcal{L}_{\mathcal{L}_X Y} f = \mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f. \qquad (6.6)$$

In the following, when writing expressions such as $\varphi_X^t(p)$, we always assume that t is close enough to 0 for this flow to be defined. With this understood, we claim that

$$f \circ \varphi_X^t = f + t g_t$$

for some smooth family of smooth real-valued functions g_t on M with $g_0 = \mathcal{L}_X f \in C^\infty(M)$.[28] This follows from the fundamental theorem of calculus: for $p \in M$ and $t \in \mathbb{R}$ close to 0, we write

$$\begin{aligned} f(\varphi_X^t(p)) - f(p) &= \int_0^1 \frac{d}{ds} f(\varphi_X^{st}(p))\, ds = \int_0^1 df\left(\partial_s \varphi_X^{st}(p)\right) ds \\ &= \int_0^1 df\left(tX(\varphi_X^{st}(p))\right) ds = t\int_0^1 df\left(X(\varphi_X^{st}(p))\right) ds, \end{aligned}$$

define $g_t(p)$ to be the integral on the right, and compute

$$g_0(p) = \int_0^1 df\left(X(\varphi_X^0(p))\right) ds = \int_0^1 df(X(p))\, ds = df(X(p)) = \mathcal{L}_X f(p),$$

proving the claim. Using this formula, we find

$$\begin{aligned} df\left([(\varphi_X^t)^* Y](p)\right) &= df\left(T\varphi_X^{-t}(Y(\varphi_X^t(p)))\right) = d(f \circ \varphi_X^{-t})(Y(\varphi_X^t(p))) \\ &= d(f - tg_t)\left(Y(\varphi_X^t(p))\right) = df\left(Y(\varphi_X^t(p))\right) - t\, dg_t\left(Y(\varphi_X^t(p))\right) \\ &= \mathcal{L}_Y f(\varphi_X^t(p)) - t\, \mathcal{L}_Y g_t(\varphi_X^t(p)). \end{aligned}$$

If we now differentiate this relation with respect to t and set $t = 0$, the left hand side becomes $df(\mathcal{L}_X Y(p)) = \mathcal{L}_{\mathcal{L}_X Y} f(p)$, while the right hand side becomes

$$d(\mathcal{L}_Y f)(X(p)) - \mathcal{L}_Y g_0(p) = \mathcal{L}_X \mathcal{L}_Y f(p) - \mathcal{L}_Y \mathcal{L}_X f(p),$$

proving (6.6). □

REMARK 6.8. The formula $\mathcal{L}_X Y = [X, Y]$ reveals that the Lie derivative of a vector field does not quite admit the interpretation we were hoping for: if $\mathcal{L}_X Y(p)$ were merely the directional derivative of $Y \in \mathfrak{X}(M)$ at p in the direction of $X \in T_pM$, then it should only depend on Y and the specific value $X(p)$, but as we see in (6.5), $[X, Y](p)$ also depends on the *first derivatives* of X at p in coordinates, not just on its value. We will see later that a straightforward directional derivative of anything more complicated than a real-valued function cannot typically be defined without making additional choices, e.g. the definition of $\mathcal{L}_X Y(p)$ requires extending $X(p)$ to a vector field that takes that value at p, and the resulting derivative depends on that choice. We will see a different and in some sense simpler way to define directional derivatives of vector fields when we study *connections* later in the semester, but a connection is also a choice that is not canonically defined in general.

6.5. Commuting flows. We can now discuss the relationship between the Lie bracket $[X, Y]$ and the question of whether the flows of X and Y commute. To understand the statement, recall from §5.1 that for each $X \in \mathfrak{X}(M)$ and $s \in \mathbb{R}$, the flow defines a diffeomorphism

$$\varphi_X^s : \mathcal{O}_X^s \to \mathcal{O}_X^{-s}$$

between two open subsets $\mathcal{O}_X^s, \mathcal{O}_X^{-s} \subset M$, which may in general be empty, but are guaranteed to be nonempty if s is close enough to 0; in fact, we have $\mathcal{O}_X^0 = \bigcup_{s>0} \mathcal{O}_X^s = \bigcup_{s<0} \mathcal{O}_X^s = M$.

[28] Saying that g_t is a "smooth family" of functions on M means literally that the function $(t, p) \mapsto g_t(p)$ for (t, p) in some open subset of $\mathbb{R} \times M$ is smooth. A slightly subtle point here is that we do not need the function $g_t : M \to M$ to be well-defined *everywhere* on M for some $t \neq 0$; for our purposes, it will suffice if $g_t(p)$ is defined for all (t, p) in some *neighborhood* of the set $\{0\} \times M$. If M is not compact, it may happen that the domain of $(t, p) \mapsto g_t(p)$ does not contain any set of the form $\{t\} \times M$ for $t \neq 0$, but is still an open neighborhood of $\{0\} \times M$.

For another vector field $Y \in \mathfrak{X}(M)$ and another $t \in \mathbb{R}$, the composition $\varphi_Y^t \circ \varphi_X^s$ is defined on $(\varphi_X^s)^{-1}(\mathcal{O}_Y^t) \subset M$, which is also open and could be empty, but is definitely not empty if both $|s|$ and $|t|$ are sufficiently small. The domain of $\varphi_X^s \circ \varphi_Y^t$ may be a different open subset of M, but is also guaranteed to overlap the domain of $\varphi_Y^t \circ \varphi_X^s$ if $|s|$ and $|t|$ are sufficiently small; in fact for every $p \in M$, there exists ϵ such that both $\varphi_X^s \circ \varphi_Y^t(p)$ and $\varphi_Y^t \circ \varphi_X^s(p)$ are defined whenever $|s|, |t| < \epsilon$.

THEOREM 6.9. *For two smooth vector fields $X, Y \in \mathfrak{X}(M)$ on a manifold M, the following conditions are equivalent:*

(i) $[X, Y] \equiv 0$;

(ii) Suppose $p \in M$ and $s, t \in \mathbb{R}$ are such that $\varphi_X^\sigma \circ \varphi_Y^\tau(p)$ is defined for all σ between 0 and s and all τ between 0 and t. Then $\varphi_Y^\tau \circ \varphi_X^\sigma(p)$ is also defined for all such σ and τ, and it equals $\varphi_X^\sigma \circ \varphi_Y^\tau(p)$. In particular, if X and Y both have global flows, then they define commuting diffeomorphisms

$$\varphi_X^s \circ \varphi_Y^t = \varphi_Y^t \circ \varphi_X^s \in \operatorname{Diff}(M)$$

for all $s, t \in \mathbb{R}$.

PROOF. We prove first that (ii) $\Rightarrow$ (i), so suppose X and Y are two vector fields whose flows commute in the sense described in the statement. For each $p \in M$, one can find a neighborhood $\mathcal{U} \subset \mathbb{R}^2$ of $(0, 0)$ small enough so that the smooth map

$$\alpha : \mathcal{U} \to M : (s, t) \mapsto \varphi_X^s \circ \varphi_Y^t(p) = \varphi_Y^t \circ \varphi_X^s(p)$$

is well-defined via either of the compositions on the right hand side. This map satisfies $\partial_s \alpha(s, t) = X(\alpha(s, t))$ and $\partial_t \alpha(s, t) = Y(\alpha(s, t))$, where the proof of the first identity requires the first version of the composition, and the second requires the second. Given $f \in C^\infty(M)$, we now define $g := f \circ \alpha : \mathcal{U} \to \mathbb{R}$ and observe that

$$\mathcal{L}_X f(\alpha(s, t)) = \partial_s g(s, t) \qquad \text{and} \qquad \mathcal{L}_Y f(\alpha(s, t)) = \partial_t g(s, t),$$

and similarly,

$$\mathcal{L}_X \mathcal{L}_Y f(\alpha(s, t)) = \partial_s \partial_t g(s, t) = \partial_t \partial_s g(s, t) = \mathcal{L}_Y \mathcal{L}_X f(\alpha(s, t)).$$

This proves in particular that $(\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X) f(p) = 0$, hence $[X, Y](p) = 0$ for all $p \in M$.

To prove (i) $\Rightarrow$ (ii), assume $[X, Y] \equiv 0$, and fix $p \in M$ and $s, t \in \mathbb{R}$ satisfying the condition specified in (ii). Then for each σ in the interval between 0 and s, φ_X^σ defines a diffeomorphism

$$M \overset{\text{open}}{\supset} \mathcal{O}_X^\sigma \xrightarrow{\varphi_X^\sigma} \mathcal{O}_X^{-\sigma} \overset{\text{open}}{\subset} M$$

whose domain and target satisfy $\mathcal{O}_X^\sigma \supset \mathcal{O}_X^s$ and $\mathcal{O}_X^{-\sigma} \supset \mathcal{O}_X^{-s}$ respectively, and moreover, the flow line $\gamma(\tau) := \varphi_Y^\tau(p)$ exists and has image in $\mathcal{O}_X^s$ for τ in the interval between 0 and t. The main step in the proof will be to show that for every σ between 0 and s, the pullback of the vector field Y from $\mathcal{O}_X^{-\sigma}$ to $\mathcal{O}_X^\sigma$ via φ_X^σ matches Y itself on $\mathcal{O}_X^s$, i.e.

$$Y = (\varphi_X^\sigma)^* Y \quad \text{on} \quad \mathcal{O}_X^s. \tag{6.7}$$

Assuming this for the moment, it then follows from Proposition 5.4 and (6.7) that the path $\tau \mapsto \varphi_X^\sigma \circ \gamma(\tau)$ for τ between 0 and t is also a flow line of Y, namely the unique one beginning at $\varphi_X^\sigma(p)$, which proves

$$\varphi_Y^\tau(\varphi_X^\sigma(p)) = \varphi_X^\sigma(\gamma(\tau)) = \varphi_X^\sigma(\varphi_Y^\tau(p)).$$

It remains only to prove (6.7). Since the statement is clearly true for $\sigma = 0$, it will suffice to prove that the derivative of the family of vector fields $(\varphi_X^\sigma)^* Y$ with respect to the parameter σ vanishes at every point on $\mathcal{O}_X^s$ for all σ between 0 and s. To see this, we use the identities $[X, Y] = \mathcal{L}_X Y = 0$ and $\varphi_X^{\sigma+\tau} = \varphi_X^\tau \circ \varphi_X^\sigma$, which gives $(\varphi_X^{\sigma+\tau})^* = (\varphi_X^\sigma)^* (\varphi_X^\tau)^*$ by Exercise 5.5.

In the following, we will only need the latter relation for values of $\tau \in \mathbb{R}$ that are arbitrarily close to 0, thus we will be free to assume that any given point in the domain of φ_X^σ is also in the domain of $\varphi_X^{\sigma+\tau}$. Working over the open set $\mathcal{O}_X^s$, we now compute,

$$\begin{aligned}\frac{d}{d\sigma}(\varphi_X^\sigma)^*Y &= \frac{d}{d\tau}(\varphi_X^{\sigma+\tau})^*Y\Big|_{\tau=0} = \frac{d}{d\tau}(\varphi_X^\sigma)^*(\varphi_X^\tau)^*Y\Big|_{\tau=0} = (\varphi_X^\sigma)^*\left(\frac{d}{d\tau}(\varphi_X^\tau)^*Y\Big|_{\tau=0}\right)\\ &= (\varphi_X^\sigma)^*(\mathcal{L}_XY) = 0.\end{aligned}$$

□

7. Tensors

It will turn out that many types of "geometric structure" on manifolds can be expressed in terms of multilinear maps on tangent and cotangent spaces, known collectively as *tensor fields.* Before beginning with the contents of this lecture, I should remind you that the Einstein summation convention (see §6.1) is in effect from now on—we are going to be needing it a lot. We will also need the following convenient notational device: for any pair of indices $i, j \in \{1, \ldots, n\}$, we define

$$\delta^{ij} = \delta_{ij} = \delta^i_j := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

The choice of whether each index is an upper or lower index will depend on the context, but the meaning will always be the same. So for example, if $\mathbf{A} \in \mathrm{GL}(n, \mathbb{R})$ is a matrix with entries $A^i{}_j$, the matrix-multiplication relation $\mathbf{A}\mathbf{A}^{-1} = \mathbb{1}$ becomes

$$A^i{}_j(A^{-1})^j{}_k = \delta^i_k.$$

Here it is very important to remember that by the summation convention, the symbol "$\sum_{j=1}^n$" has been omitted from the left hand side; we chose to write the first index of $A^i{}_j$ as an upper index and the second as a lower index mainly so that this use of the summation convention would work. Here is another example that already came up in our discussion of vector fields (cf. Exercise 6.2): if $(\mathcal{U}, x)$ and $(\widetilde{\mathcal{U}}, \widetilde{x})$ are two overlapping charts on a manifold M, then at every point in $\mathcal{U} \cap \widetilde{\mathcal{U}}$, the matrices with entries $\frac{\partial \widetilde{x}^i}{\partial x^j}$ and $\frac{\partial x^i}{\partial \widetilde{x}^j}$ are inverse to each other, as they are Jacobi matrices of inverse transition maps, thus

$$\frac{\partial \widetilde{x}^i}{\partial x^j}\frac{\partial x^j}{\partial \widetilde{x}^k} = \delta^i_k.$$

Other versions of δ will sometimes arise with the indices placed in various ways in order to make the summation convention work. This symbol is known as the **Kronecker delta**, and maybe it would have been called something different if it had been invented in the age of Covid-19, but here we are.

7.1. Motivational examples. In order to motivate the idea of a tensor field on a manifold, it's best to start with a few examples that are already somewhat familiar.

7.1.1. *One-forms.* Any smooth function $f : M \to \mathbb{R}$ has a differential

$$df : TM \to \mathbb{R},$$

whose restriction to each individual tangent space T_pM is a linear map $T_pM \to \mathbb{R}$ and thus an element of the cotangent space T_p^*M. In this sense, df is analogous to a vector field, but instead of associating a tangent vector $X(p) \in T_pM$ to every point $p \in M$, it associates a cotangent vector $d_pf \in T_p^*M$, thus defining a map

$$M \to T^*M : p \mapsto d_pf.$$

In general, a map

$$\lambda : TM \to \mathbb{R}$$

whose restriction to each individual tangent space is linear is called a **1-form** on M, or sometimes also a **dual vector field** or **covector field**. For each $p \in M$, it is common to denote the restriction $\lambda|_{T_pM} : T_pM \to \mathbb{R}$ by

$$\lambda_p \in T_p^*M = \mathrm{Hom}(T_pM, \mathbb{R}),$$

hence one can equivalently view a 1-form λ as associating to each point $p \in M$ a cotangent vector $\lambda_p \in T_p^*M$. For the special case where λ is the differential of a function f, we have been writing $d_pf \in T_p^*M$ for the restriction to T_pM, but the notation $(df)_p$ would also be sensible, and is preferred by many authors.[29]

Since we have not yet endowed the cotangent bundle T^*M with a smooth structure, we need to put some thought into defining what it means for a 1-form to be "smooth". The easiest way to do this is by writing it in local coordinates. Any chart $(\mathcal{U}, x)$ on M gives rise to coordinate functions $x^i : \mathcal{U} \to \mathbb{R}$ for $i = 1, \ldots, n$, whose differentials dx^i are 1-forms on $\mathcal{U}$.

PROPOSITION 7.1. *For each $p \in \mathcal{U}$, every element $\lambda \in T_p^*M$ can be expressed as a linear combination $\lambda = \lambda_i\, d_px^i$ for unique real numbers $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. In other words, the differentials $d_px^1, \ldots, d_px^n$ form a basis of T_p^*M.*

PROOF. What's actually happening here is that $d_px^1, \ldots, d_px^n$ is the dual basis to the basis of coordinate vector fields $\partial_1, \ldots, \partial_n$ defined by the chart $(\mathcal{U}, x)$ at p; indeed, for each $i, j \in \{1, \ldots, n\}$,

$$dx^i(\partial_j) = dx^i\left(\frac{\partial}{\partial x^j}\right) = \mathcal{L}_{\frac{\partial}{\partial x^j}} x^i = \frac{\partial x^i}{\partial x^j} = \delta^i_j.$$

The coefficients λ_i are thus given by $\lambda_i = \lambda(\partial_i)$. □

The 1-forms $dx^1, \ldots, dx^n$ on $\mathcal{U}$ defined by a chart $(\mathcal{U}, x)$ are known as the **coordinate differentials**, and Proposition 7.1 implies that every 1-form λ can be written over the region $\mathcal{U}$ as

$$\lambda = \lambda_i\, dx^i,$$

where its uniquely determined **component** functions $\lambda_i : \mathcal{U} \to \mathbb{R}$ are given by

$$\lambda_i(p) := \lambda\left(\frac{\partial}{\partial x^i}(p)\right), \qquad p \in \mathcal{U}.$$

For example, the component functions of the differential df are precisely the partial derivatives of f, namely $(df)_i = df(\partial_i) = \partial_i f : \mathcal{U} \to \mathbb{R}$, giving rise to the formula

$$df = \partial_i f\, dx^i \qquad \text{on } \mathcal{U},$$

which was understood for at least two centuries in terms of "infinitessimal quantities" before it was given a mathematically rigorous meaning in terms of 1-forms.

REMARK 7.2. Notice that while components of vector fields are written with upper indices, components of 1-forms get lower indices. This is necessary in order for the summation convention to work properly, since coordinate differentials come with upper indices.

[29] Or if one prefers to think of df as a function $M \to T^*M$, one can write $df(p)$ instead of d_pf or $(df)_p$. I have done that in some of my research papers, but will avoid it in these notes for the sake of consistency, as we have defined df as a function $TM \to \mathbb{R}$ rather than $M \to T^*M$.

EXERCISE 7.3. Suppose $(\mathcal{U}, x)$ and $(\widetilde{\mathcal{U}}, \widetilde{x})$ are two smooth charts with $\mathcal{U} \cap \widetilde{\mathcal{U}} \neq \emptyset$, so any 1-form λ can be written as both $\lambda_i \, dx^i$ and $\widetilde{\lambda}_i \, d\widetilde{x}^i$ in the overlap region. Prove the following coordinate transformation formulas on $\mathcal{U} \cap \widetilde{\mathcal{U}}$, analogous to the formulas (6.3) and (6.4) for vector fields:

$$dx^i = \frac{\partial x^i}{\partial \widetilde{x}^j} \, d\widetilde{x}^j \qquad \text{and} \qquad \widetilde{\lambda}_i = \lambda_j \frac{\partial x^j}{\partial \widetilde{x}^i}. \tag{7.1}$$

The formula (7.1) shows that if a 1-form has smooth component functions with respect to any given chart, its component functions in any other chart defined on the same domain will also be smooth, due to the fact that transition maps (and therefore also their derivatives $\frac{\partial x^i}{\partial \widetilde{x}^j}$) are smooth. The following definition therefore makes sense.

DEFINITION 7.4. A 1-form on M is said to be **smooth** if and only if its component functions with respect to every chart are smooth. The set of all smooth 1-forms on M forms a vector space, which we denote by

$$\Omega^1(M) := \{\text{smooth 1-forms on } M\}.$$

EXERCISE 7.5. Show that a 1-form λ on M is smooth if and only if the function $M \to \mathbb{R} : p \mapsto \lambda(X(p))$ is smooth for every smooth vector field $X \in \mathfrak{X}(M)$.

From now on, we will assume that all 1-forms we consider are smooth unless stated otherwise.

7.1.2. *Vector fields.* Recall that every finite-dimensional vector space V is naturally isomorphic to the dual of its dual space, with a canonical isomorphism $\Phi : V \to V^{**}$ given by

$$\Phi(v)\lambda := \lambda(v).$$

If we choose to, we can therefore also think of every tangent space T_pM as a dual space, namely $(T_p^*M)^*$, meaning that every vector field $X \in \mathfrak{X}(M)$ can equivalently be viewed as associating to each $p \in M$ a linear map $\tau_p : T_p^*M \to \mathbb{R}$, defined by $\tau_p(\lambda) := \lambda(X(p))$. I'm sure you can imagine why we didn't define vector fields this way in the first place, but we could have done so if we'd wanted to. From this perspective, the notion of smoothness for a vector field can also be characterized analogously to Exercise 7.5:

EXERCISE 7.6. Show that a vector field X on M is smooth if and only if the function $M \to \mathbb{R} : p \mapsto \lambda(X(p))$ is smooth for every smooth 1-form $\lambda \in \Omega^1(M)$.

7.1.3. *Riemannian metrics.* A Riemannian metric g on a manifold M associates to every point $p \in M$ an inner product g_p on T_pM, so in particular, g_p is a bilinear map

$$g_p : T_pM \times T_pM \to \mathbb{R}$$

that is also symmetric and positive-definite. We can think of g itself as a function

$$g : TM \oplus TM \to \mathbb{R},$$

where $TM \oplus TM := \bigcup_{p \in M} (T_pM \times T_pM)$. As a provisional notion of smoothness for Riemannian metrics, we can define g to be **smooth** if and only if the function

$$M \to \mathbb{R} : p \mapsto g(X(p), Y(p))$$

is smooth for every pair of smooth vector fields $X, Y \in \mathfrak{X}(M)$. Under this condition, g is an example of something we will shortly define as a "smooth covariant tensor field of rank 2" on M.

7.1.4. *Almost complex structures.* Here is an example you may not have heard of before. One can make any $2n$-dimensional real vector space V into an n-dimensional complex vector space by choosing a linear map $J : V \to V$ with $J^2 = -\mathbb{1}$ and defining complex scalar multiplication on V by $(a + ib)v := av + bJv$. Such a linear map J is therefore called a **complex structure** on V. It is sometimes useful to introduce such a structure on the tangent spaces of an even-dimensional manifold M. An **almost complex structure** (*fast komplexe Struktur*) on M is a map

$$J : TM \to TM$$

whose restriction to each individual tangent space is a complex structure $J_p : T_pM \to T_pM$. We can define J to be **smooth** if and only if the vector field $p \mapsto JX(p)$ is smooth for all smooth vector fields $X \in \mathfrak{X}(M)$. The following lemma gives an alternative algebraic way of understanding what an almost complex structure is.

Lemma 7.7. *For a finite-dimensional real vector space V, let $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ denote the vector space of all linear maps $V \to V$, $V^* = \mathrm{Hom}(V, \mathbb{R})$ the dual space of V, and $\mathrm{Hom}(V^* \otimes V, \mathbb{R})$ the vector space of all bilinear maps $V^* \times V \to \mathbb{R}$. There exists a canonical isomorphism*

$$\Phi : \mathrm{End}(V) \to \mathrm{Hom}(V^* \otimes V, \mathbb{R}), \qquad \Phi(A)(\lambda, v) := \lambda(Av).$$

Proof. It is easy to check that Φ is a linear injection, and if $\dim V = n$, then $\dim \mathrm{End}(V) = \dim \mathrm{Hom}(V^* \otimes V, \mathbb{R}) = n^2$, thus Φ is also surjective. □

For an almost complex structure J on M, Lemma 7.7 allows us to view $J_p : T_pM \to T_pM$ equivalently as a bilinear map $T_p^*M \times T_pM \to \mathbb{R}$, and from this perspective, one can check that J is smooth (according to our previous definition) if and only if the function $M \to \mathbb{R} : p \mapsto J(\lambda_p, X(p))$ is smooth for all choices of smooth vector field $X \in \mathfrak{X}(M)$ and smooth 1-form $\lambda \in \Omega^1(M)$.

7.2. Tensor fields in general. We now describe a more general notion that encompasses all of the examples in §7.1 as special cases.

Recall that for vector spaces $V_1, \ldots, V_n$ and W, a map

$$T : V_1 \times \ldots \times V_n \to W$$

is called **multilinear** if it is linear with respect to each variable individually, i.e. for every $i = 1, \ldots, n$ and every fixed tuple of vectors $v_j \in V_j$ for $j = 1, \ldots, i-1, i+1, \ldots, n$, the map

$$V_i \to W : v_i \mapsto T(v_1, \ldots, v_n)$$

is linear. Observe that the space of all multilinear maps $V_1 \times \ldots \times V_n \to W$ is naturally also a finite-dimensional vector space. We will sometimes denote it by[30]

$$\mathrm{Hom}(V_1 \otimes \ldots \otimes V_n, W).$$

Definition 7.8. For integers $k, \ell \geqslant 0$ with $k + \ell > 0$ and a finite-dimensional real vector space V, we will denote by V_ℓ^k the vector space of multilinear maps

$$\underbrace{V^* \times \ldots \times V^*}_{k} \times \underbrace{V \times \ldots \times V}_{\ell} \to \mathbb{R},$$

where V^* as usual denotes the dual space $\mathrm{Hom}(V, \mathbb{R})$. In the case $k = \ell = 0$, we define $V_0^0 = \mathbb{R}$.

[30] We will not make use of the abstract algebraic notion of the tensor product of vector spaces in this lecture, but readers already familiar with that notion may want to pause and consider why our definition of the symbol "$\mathrm{Hom}(V_1 \otimes \ldots \otimes V_n, W)$" is equivalent to the one they've seen before. It is important that we are explicitly assuming all vector spaces to be finite dimensional in this discussion; if we did not assume this, then some more serious digressions into the meaning of the symbol "$\otimes$" would be necessary.

REMARK 7.9. To motivate the convention $V_0^0 = \mathbb{R}$, you can imagine perhaps that a "real-valued multilinear function of *zero* variables" is the same thing as a real number. If that doesn't convince you, the convention will at least begin to seem more natural when we discuss tensor products (cf. Remark 7.19).

DEFINITION 7.10. For a smooth manifold M and integers $k, \ell \geqslant 0$, a **tensor field** (*Tensorfeld*) S of type (k, ℓ) associates to each point $p \in M$ an element

$$S_p \in (T_pM)^k_\ell.$$

If $k + \ell > 0$, then the tensor field S is said to be **smooth** if and only if the function $M \to \mathbb{R} : p \mapsto S_p(\lambda^1_p, \ldots, \lambda^k_p, X_1(p), \ldots, X_\ell(p))$ is smooth for every tuple of smooth vector fields $X_1, \ldots, X_\ell \in \mathfrak{X}(M)$ and smooth 1-forms $\lambda^1, \ldots, \lambda^k \in \Omega^1(M)$. We will denote the vector space of smooth tensor fields by

$$\Gamma(T^k_\ell M) := \{\text{smooth tensor fields of type } (k, \ell)\}.$$

For $k = \ell = 0$, a tensor field is just a real-valued function on M, so we define $\Gamma(T^0_0 M) := C^\infty(M)$.

The **support** (*Träger*) of a tensor field $S \in \Gamma(T^k_\ell M)$ is defined as the closure in M of the set $\{p \in M \mid S_p \neq 0\}$.

EXAMPLE 7.11. A smooth 1-form is equivalently a smooth tensor field of type $(0, 1)$:

$$\Omega^1(M) = \Gamma(T^0_1 M).$$

Just as 1-forms $\lambda \in \Omega^1(M)$ are regarded as functions $TM \to \mathbb{R}$, it will often be useful to regard a tensor field $S \in \Gamma(T^k_\ell M)$ in the case $k + \ell > 0$ as a function

$$S : T^*M^{\oplus k} \oplus TM^{\oplus \ell} \to \mathbb{R},$$

where we introduce the notation

$$T^*M^{\oplus k} \oplus TM^{\oplus \ell} := \bigcup_{p \in M} \Bigg(\underbrace{T^*_pM \times \ldots \times T^*_pM}_{k} \times \underbrace{T_pM \times \ldots \times T_pM}_{\ell} \Bigg).$$

The key property of S is then that its restriction S_p to $T^*_pM \times \ldots \times T^*_pM \times T_pM \times \ldots \times T_pM \subset T^*M^{\oplus k} \oplus TM^{\oplus \ell}$ for each $p \in M$ is a multilinear map.

In the setting of smooth manifolds, the term "tensor field" is often abbreviated simply as **tensor**. The terminology for tensors of type (k, ℓ) can also vary among different sources, e.g. one sometimes says that a tensor $S \in \Gamma(T^k_\ell M)$ is **contravariant of rank** k and **covariant of rank** ℓ. The latter terminology is especially favored among physicists.

EXAMPLE 7.12. Under the canonical isomorphism identifying each tangent space T_pM with $\mathrm{Hom}(T^*_pM, \mathbb{R})$, a smooth vector field becomes the same thing as a smooth tensor field of type $(1, 0)$, hence

$$\mathfrak{X}(M) = \Gamma(T^1_0 M).$$

Here the function $T^*M \to \mathbb{R}$ corresponding to a given vector field $X \in \mathfrak{X}(M)$ sends $\lambda \in T^*_pM$ to $\lambda(X(p))$.

EXAMPLE 7.13. Every Riemannian metric (see §7.1.3) is an example of a tensor field of type $(0, 2)$.

EXAMPLE 7.14. Every almost complex structure (see §7.1.4) is an example of a tensor field of type $(1, 1)$.

EXERCISE 7.15. Generalize Lemma 7.7 to show the following: for any finite-dimensional real vector spaces $V_1, \ldots, V_n, W$, there exists a canonical isomorphism

$$\operatorname{Hom}(V_1 \otimes \ldots \otimes V_n, W) \xrightarrow{\Phi} \operatorname{Hom}(W^* \otimes V_1 \otimes \ldots \otimes V_n, \mathbb{R}),$$
$$\Phi(A)(\lambda, v_1, \ldots, v_n) := \lambda(A(v_1, \ldots, v_n)).$$

EXAMPLE 7.16. For arbitrary integers $\ell \geqslant 1$, Exercise 7.15 identifies any tensor field S of type $(1, \ell)$ with a map

$$\bigcup_{p \in M} \Big(\underbrace{T_pM \times \ldots \times T_pM}_{\ell} \Big) =: TM^{\oplus \ell} \xrightarrow{\widehat{S}} TM$$

whose restriction $\widehat{S}_p$ to $T_pM \times \ldots \times T_pM$ for each $p \in M$ is a multilinear map $T_pM \times \ldots \times T_pM \to T_pM$. The precise correspondence between S and $\widehat{S}$ is given by

$$S(\lambda, X_1, \ldots, X_\ell) = \lambda\big(\widehat{S}(X_1, \ldots, X_\ell)\big),$$

and it is straightforward to show that S is smooth if and only if $\widehat{S}(X_1, \ldots, X_\ell)$ defines a smooth vector field for all choices of smooth vector fields $X_1, \ldots, X_\ell \in \mathfrak{X}(M)$. The case $\ell = 0$ also fits into this picture if one adopts the perspective that a "T_pM-valued function of zero variables" just means an element of T_pM: this reproduces the observation in Example 7.12 that tensor fields of type $(1, 0)$ are equivalent to vector fields.

REMARK 7.17. The alternative perspective on tensors of type $(1, \ell)$ in Example 7.16 will generally be quite useful, and from now on we will typically use the same notation for the objects that are called S and $\widehat{S}$ in that example. We have already adopted this convention in our discussion of vector fields and almost complex structures as tensors of type $(1, 0)$ and $(1, 1)$ respectively.

DEFINITION 7.18. For $S \in \Gamma(T^k_\ell M)$ and $T \in \Gamma(T^r_s M)$, the **tensor product** (*Tensorprodukt*) of S and T is the tensor field $S \otimes T \in \Gamma(T^{k+r}_{\ell+s} M)$ defined at each point $p \in M$ by

$$(S \otimes T)_p(\lambda^1, \ldots, \lambda^k, \mu^1, \ldots, \mu^r, X_1, \ldots, X_\ell, Y_1, \ldots, Y_s) := S_p(\lambda^1, \ldots, \lambda^k, X_1, \ldots, X_\ell) \cdot T_p(\mu^1, \ldots, \mu^r, Y_1, \ldots, Y_s).$$

REMARK 7.19. For $f \in C^\infty(M) = \Gamma(T^0_0 M)$, the tensor product of f with $S \in \Gamma(T^k_\ell M)$ is just the ordinary point-wise product of S with a scalar-valued function, i.e. $(f \otimes S)_p = (S \otimes f)_p = f(p) S_p$.

7.3. Coordinate representations. We've seen that a chart $(\mathcal{U}, x)$ on M gives rise to coordinate vector fields $\partial_1, \ldots, \partial_n \in \mathfrak{X}(\mathcal{U})$ and coordinate differentials $dx^1, \ldots, dx^n \in \Omega^1(\mathcal{U})$ which define bases of T_pM and T_p^*M respectively at each point $p \in \mathcal{U}$. Regarding vector fields as tensors of type $(1, 0)$, it turns out that a natural basis of $(T_pM)^k_\ell$ can then be constructed by taking all possible tensor products of k coordinate vector fields with ℓ coordinate differentials. Indeed:

PROPOSITION 7.20. *Given a chart $(\mathcal{U}, x)$ on an n-manifold M, every tensor field S of type (k, ℓ) can be written uniquely over $\mathcal{U}$ as*

$$S = S^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_k}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_\ell}, \tag{7.2}$$

*where the $n^{k+\ell}$ **component** functions $S^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell} : \mathcal{U} \to \mathbb{R}$ are given by*

$$S^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell} := S(dx^{i_1}, \ldots, dx^{i_k}, \partial_{j_1}, \ldots, \partial_{j_\ell}).$$

Remark 7.21. Writing down (7.2) without the Einstein summation convention would have required inserting the symbols

$$\sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \sum_{j_1=1}^{n} \cdots \sum_{j_\ell=1}^{n}$$

just to the right of the equal sign, so the right hand side is actually a sum of $n^{k+\ell}$ terms.

Proof of Proposition 7.20. Any ℓ vector fields can be written over $\mathcal{U}$ as $X_a = X_a^i\,\partial_i$ for $a = 1, \dots, \ell$ with unique component functions $X_a^i : \mathcal{U} \to \mathbb{R}$, and similarly, any k 1-forms can be written as $\lambda^b = \lambda^b_j\,dx^j$ with unique components $\lambda^b_j : \mathcal{U} \to \mathbb{R}$. By multilinearity, we then have

$$\begin{aligned} S(\lambda^1, \dots, \lambda^k, X_1, \dots, X_\ell) &= S(\lambda^1_{i_1}\,dx^{i_1}, \dots, \lambda^k_{i_k}\,dx^{i_k}, X_1^{j_1}\,\partial_{j_1}, \dots, X_\ell^{j_\ell}\,\partial_{j_\ell}) \\ &= S^{i_1 \dots i_k}{}_{j_1 \dots j_\ell}\,\lambda^1_{i_1} \dots \lambda^k_{i_k} X_1^{j_1} \dots X_\ell^{j_\ell}. \end{aligned} \tag{7.3}$$

It is straightforward to check that the tensor field on the right hand side of (7.2) gives the same result when evaluated on the same tuple of vector fields and 1-forms. □

Exercise 7.22. Show that a tensor field of type (k, ℓ) is smooth if and only if for every smooth chart, the corresponding component functions are all smooth.

Exercise 7.23. Show that in local coordinates, the components of two tensor fields $S \in \Gamma(T^k_\ell M)$, $T \in \Gamma(T^r_s M)$ and their tensor product $S \otimes T \in \Gamma(T^{k+r}_{\ell+s} M)$ are related by

$$(S \otimes T)^{i_1 \dots i_k a_1 \dots a_r}{}_{j_1 \dots j_\ell b_1 \dots b_s} = S^{i_1 \dots i_k}{}_{j_1 \dots j_\ell}\, T^{a_1 \dots a_r}{}_{b_1 \dots b_s}.$$

Exercise 7.24. Suppose $(\mathcal{U}, x)$ and $(\widetilde{\mathcal{U}}, \widetilde{x})$ are two smooth charts with $\mathcal{U} \cap \widetilde{\mathcal{U}} \ne \varnothing$, and denote the component functions of a tensor field $S \in \Gamma(T^k_\ell M)$ with respect to each chart by $S^{i_1 \dots i_k}{}_{j_1 \dots j_\ell}$ and $\widetilde{S}^{i_1 \dots i_k}{}_{j_1 \dots j_\ell}$ respectively. Prove that on the overlap region $\mathcal{U} \cap \widetilde{\mathcal{U}}$,

$$\widetilde{S}^{i_1 \dots i_k}{}_{j_1 \dots j_\ell} = \frac{\partial \widetilde{x}^{i_1}}{\partial x^{a_1}} \cdots \frac{\partial \widetilde{x}^{i_k}}{\partial x^{a_k}} S^{a_1 \dots a_k}{}_{b_1 \dots b_\ell} \frac{\partial x^{b_1}}{\partial \widetilde{x}^{j_1}} \cdots \frac{\partial x^{b_\ell}}{\partial \widetilde{x}^{j_\ell}}. \tag{7.4}$$

Hint: Use (6.3) and (7.1).

Remark 7.25. We have been writing all tensor fields so far as functions that take covectors $\lambda^1, \dots, \lambda^k$ followed by vectors $X_1, \dots, X_\ell$, but in some circumstances, one may want to be more flexible with the ordering, so that e.g. a tensor of type $(1, 2)$ could be written as a multilinear function

$$TM \oplus T^*M \oplus TM \to \mathbb{R} : (X, \lambda, Y) \mapsto S(X, \lambda, Y).$$

The component functions of such a tensor would then be written as $S_i{}^j{}_k$, with evaluation on $X = X^i\,\partial_i$, $\lambda = \lambda^j\,dx^j$ and $Y = Y^k\,\partial_k$ defined by the rule

$$S(X, \lambda, Y) = S_i{}^j{}_k X^i \lambda_j Y^k.$$

Example 7.26. Suppose $J : TM \to TM$ is an almost complex structure, so $J_p : T_pM \to T_pM$ is a linear map satisfying $J_p^2 = -\mathbb{1}$ for every $p \in M$. As we've seen, J can be regarded as a tensor field of type $(1, 1)$ and thus defines a function $T^*M \oplus TM \to \mathbb{R}$, with component functions with respect to a chart $(\mathcal{U}, x)$ written as

$$J^i{}_j = J(dx^i, \partial_j) := dx^i(J\partial_j), \qquad i, j \in \{1, \dots, n\}.$$

In this line, the second expression views J_p as a bilinear map $T_p^*M \times T_pM \to \mathbb{R}$, while the third views it as a linear map $T_pM \to T_pM$. This means that for two tangent vectors $X = X^i\,\partial_i$ and $Y = Y^i\,\partial_i$ at a point $p \in \mathcal{U}$, we have

$$JX = Y \qquad \Longleftrightarrow \qquad Y^i = dx^i(Y) = dx^i(JX) = dx^i(J(X^j\,\partial_j)) = X^j\,dx^i(J\partial_j) = J^i{}_j X^j,$$

so in other words, the linear map $J_p : T_pM \to T_pM$ is represented in coordinates by matrix-vector multiplication: the n-by-n matrix with entries $J^i{}_j$ gets multiplied by the n-dimensional row vector with entries X^j to produce the row vector with entries $(JX)^i$. The condition $J^2 = -\mathbb{1}$ can thus be expressed in local coordinates on $\mathcal{U}$ as

$$J^i{}_j J^j{}_k \equiv -\delta^i_k \qquad \text{on } \mathcal{U}.$$

From this perspective, the transformation formula (7.4) also ends up looking like something familiar from linear algebra: the component functions $J^i{}_j$ and $\widetilde{J}^i{}_j$ for two overlapping charts $(\mathcal{U}, x)$ and $(\widetilde{\mathcal{U}}, \widetilde{x})$ are related by

$$\widetilde{J}^i{}_j = \frac{\partial \widetilde{x}^i}{\partial x^k} J^k{}_\ell \frac{\partial x^\ell}{\partial \widetilde{x}^j}.$$

In terms of matrices, this just says

$$\widetilde{\mathbf{J}} = \left(\frac{\partial \widetilde{x}}{\partial x}\right) \mathbf{J} \left(\frac{\partial \widetilde{x}}{\partial x}\right)^{-1},$$

where $\mathbf{J}$ and $\widetilde{\mathbf{J}}$ denote the n-by-n matrices with entries $J^i{}_j$ and $\widetilde{J}^i{}_j$ respectively, while $\frac{\partial \widetilde{x}}{\partial x}$ is the n-by-n Jacobian matrix with entries $\frac{\partial \widetilde{x}^i}{\partial x^j}$.

8. Derivatives of tensors and differential forms

We motivate this lecture with the following question: for a smooth tensor field $S \in \Gamma(T^k_\ell M)$, can one define a "directional derivative" of S at a point $p \in M$ in the direction $X \in T_pM$? We considered this question for the special case of vector fields $Y \in \mathfrak{X}(M) = \Gamma(T^1_0 M)$ in §6.4, and the answer we came up with there was not entirely satisfactory: a vector field Y can be differentiated with respect to another vector field X, producing the Lie derivative $\mathcal{L}_X Y \in \mathfrak{X}(M)$, but $\mathcal{L}_X Y(p)$ depends on X as a vector field, not just on the value $X(p)$ (see Remark 6.8). Naively, one might hope for instance that if $S \in \Gamma(T^k_\ell M)$ has components $S^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell}$ with respect to some chart $(\mathcal{U}, x)$, then one could define a tensor "dS" of type $(k, \ell + 1)$ whose components are

$$\text{“}(dS)^{i_1 \ldots i_k}{}_{j_0 \ldots j_\ell} = \partial_{j_0} S^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell}\text{”}, \tag{8.1}$$

so that for any $p \in M$ and $X \in T_pM$, the multilinear map $(dS)(\ldots, X, \ldots) : (T^*_p M)^{\times k} \times (T_pM)^{\times \ell} \to \mathbb{R}$ could be interpreted as the derivative of S in the direction X. But I put that expression in quotation marks because, indeed, it doesn't work: outside of the special case $k = \ell = 0$ where the objects we are differentiating are just real-valued functions, one cannot define from $S \in \Gamma(T^k_\ell M)$ any tensor field $dS \in \Gamma(T^k_{\ell+1} M)$ whose components are given in *all* choices of local coordinates by (8.1). (Exercise 8.1(b) below asks you to prove this in the case $(k, \ell) = (0, 1)$.) In other words, the formula (8.1) is not *coordinate invariant*.

Before discussing directional derivatives further, we should talk about a sticky issue that arose in the previous paragraph: what *practical* methods do we have for writing down the definition of a tensor field? What we attempted above could be called the *physicists' method*: it starts by choosing a chart $(\mathcal{U}, x)$ and writing down a formula for the component functions of the tensor with respect to those local coordinates. That is fine if one only needs a tensor field defined on the subset $\mathcal{U} \subset M$, but the hope of course is that the formula we write down might be valid in *arbitrary* local coordinates, in which case it gives a well-defined tensor field everywhere on M. The important step is therefore to check, using the transformation formula (7.4), that the definition we've written is coordinate invariant, and that is what fails in the case of (8.1). On the other hand, sometimes it succeeds, for instance:

Exercise 8.1. Prove:

(a) For any $\lambda \in \Gamma(T_1^0 M)$, there exists a tensor field $S \in \Gamma(T_2^0 M)$ whose components S_{ij} with respect to arbitrary charts $(\mathcal{U}, x)$ are related to the corresponding components λ_i of λ by

$$S_{ij} = \partial_i \lambda_j - \partial_j \lambda_i.$$

(b) For general choices of λ, one cannot similarly define $S \in \Gamma(T_2^0 M)$ so that its relation to λ in arbitrary local coordinates is $S_{ij} = \partial_i \lambda_j$.

Physicists like to summarize the result of Exercise 8.1(a) by saying that the expression $\partial_i \lambda_j - \partial_j \lambda_i$ "defines a tensor" of type $(0,2)$. In fact, many textbooks on general relativity give a definition of tensors that is cosmetically quite different from ours: without mentioning multilinear maps, they define a tensor S of type (k,ℓ) as an association to each chart $(\mathcal{U}, x)$ of a collection of real-valued functions $S^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell} : \mathcal{U} \to \mathbb{R}$ that satisfy the transformation formula (7.4). There are good theoretical reasons why mathematicians do not usually give that as the definition of a tensor field, and contrary to what many physicists may tell you, it is also not true that defining a tensor or computing something from it always requires choosing local coordinates.

8.1. C^∞-linearity. Here is a trick for writing down tensor fields that mathematicians tend to prefer, because it does not require local coordinates. For example, let us regard a tensor field S of type $(1,\ell)$ as associating to each point $p \in M$ an ℓ-fold multilinear map $S_p : T_p M \times \ldots \times T_p M \to T_p M$, as described in Example 7.16. It therefore also defines a multilinear map

$$S : \underbrace{\mathfrak{X}(M) \times \ldots \times \mathfrak{X}(M)}_{\ell} \to \mathfrak{X}(M), \tag{8.2}$$

by interpreting $S(X_1, \ldots, X_\ell)$ for any tuple of smooth vector fields $X_1, \ldots, X_\ell$ as the vector field

$$p \mapsto S_p(X_1(p), \ldots, X_\ell(p)).$$

We already know one important concrete example of multilinear map of this type: the Lie bracket is a bilinear map

$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M).$$

But does the Lie bracket therefore define a tensor field of type $(1,2)$? It would be surprising if this were true, because being a tensor field would imply that the value $[X,Y](p)$ for each $p \in M$ depends only on the values $X(p)$ and $Y(p)$, whereas we saw in Exercise 6.2 that in local coordinates, $[X,Y](p)$ also depends on the first derivatives of X and Y at p. An easy way to make this intuition more precise is via the following observation: if S is a tensor field, then the map in (8.2) is not just multilinear, it also satisfies

$$S(X_1, \ldots, X_{j-1}, f X_j, X_{j+1}, \ldots, X_\ell) = f S(X_1, \ldots, X_\ell) \quad \text{for all } f \in C^\infty(M) \tag{8.3}$$

for every $j = 1, \ldots, \ell$. The key point here is that the function f does not need to be constant, so this is a much stronger statement than just saying that (8.2) respects scalar multiplication (as every multilinear map must). A multilinear map on the space of vector fields is said to be **C^∞-linear** in its jth argument if it satisfies (8.3). In general, the notion of C^∞-linearity can be defined for multilinear maps between any vector spaces on which there is a natural notion of multiplication by smooth functions[31], e.g. we had $\mathfrak{X}(M)$ in the above example because the product of a smooth vector field with a smooth function is also a smooth vector field, but for similar reasons, one could just as well work with $\Omega^1(M)$, the other spaces of smooth tensor fields $\Gamma(T_\ell^k M)$, or $C^\infty(M)$ itself. From this perspective, the obvious reason why the Lie bracket does not define a tensor field is that it is not C^∞-linear: according to Exercise 6.4, it satisfies

$$[fX, Y] = f[X,Y] - (\mathcal{L}_Y f) X, \qquad [X, fY] = f[X,Y] + (\mathcal{L}_X f) Y,$$

[31] in other words, spaces that are naturally *modules* over $C^\infty(M)$

for $f \in C^\infty(M)$, which is not the desired relation except in the special case where f is constant.

It will be exceedingly useful to observe that C^∞-linearity is not only necessary for a multilinear map on vector fields or 1-forms to define a tensor field—it is also sufficient.

PROPOSITION 8.2. *For a multilinear map*

$$S : \underbrace{\Omega^1(M) \times \ldots \times \Omega^1(M)}_{k} \times \underbrace{\mathfrak{X}(M) \times \ldots \times \mathfrak{X}(M)}_{\ell} \to C^\infty(M)$$

that is C^∞-linear in every argument, there exists a unique tensor field $\hat{S} \in \Gamma(T^k_\ell M)$ such that for every $p \in M$, $X_1, \ldots, X_\ell \in \mathfrak{X}(M)$ and $\lambda^1, \ldots, \lambda^k \in \Omega^1(M)$,

$$\hat{S}(\lambda^1_p, \ldots, \lambda^k_p, X_1(p), \ldots, X_\ell(p)) = S(\lambda^1, \ldots, \lambda^k, X_1, \ldots, X_\ell)(p).$$

Before proving the theorem, let us observe that it can be adapted easily for the slightly different situation in (8.2), where our multilinear map takes values in $\mathfrak{X}(M)$ instead of $C^\infty(M)$:

EXERCISE 8.3. Deduce from Proposition 8.2 that for any multilinear map

$$S : \underbrace{\mathfrak{X}(M) \times \ldots \times \mathfrak{X}(M)}_{\ell} \to \mathfrak{X}(M)$$

that is C^∞-linear in every argument, there exists a unique tensor field $\hat{S} \in \Gamma(T^1_\ell M)$ such that for every $p \in M$, $X_1, \ldots, X_\ell \in \mathfrak{X}(M)$, the multilinear map $\hat{S}_p : T_pM \times \ldots \times T_pM \to T_pM$ satisfies

$$\hat{S}(X_1(p), \ldots, X_\ell(p)) = S(X_1, \ldots, X_\ell)(p).$$

PROOF OF PROPOSITION 8.2. Let us consider only the case $\ell = 1$ and $k = 0$, as there is no substantial difference in the general case beyond requiring more complicated notation. We therefore assume $\Lambda : \mathfrak{X}(M) \to C^\infty(M)$ is a linear map satisfying $\Lambda(fX) = f\Lambda(X)$ for all $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$, and we need to find a smooth 1-form $\lambda \in \Omega^1(M)$ such that $\lambda(X(p)) = \Lambda(X)(p)$ for all $p \in M$ and $X \in \mathfrak{X}(M)$. The uniqueness of λ is clear, since every tangent vector at a point $p \in M$ can be the value at that point of a smooth vector field (just write it down in local coordinates, multiply by a smooth cutoff function and extend outside of the coordinate neighborhood as 0).

To prove existence, it suffices to show that for any point $p \in M$, the value of $\Lambda(X)(p)$ is completely determined by $X(p)$ and does not otherwise depend on the choice of vector field X having this particular value at p. This will follow from linearity after proving two claims:

***Claim 1**: If $X \in \mathfrak{X}(M)$ vanishes in a neighborhood of p, then $\Lambda(X)(p) = 0$.*

Indeed, if $\mathcal{U} \subset M$ is an open neighborhood on which X vanishes, choose a smooth function $\beta : M \to [0, 1]$ with compact support in $\mathcal{U}$ satisfying $\beta(p) = 1$. Then $\beta X \equiv 0$, thus by C^∞-linearity,

$$0 = \Lambda(\beta X) = \beta\Lambda(X) \in C^\infty(M),$$

implying in particular that $\Lambda(X)(p) = \beta(p)\Lambda(X)(p) = 0$.

***Claim 2**: If $X \in \mathfrak{X}(M)$ satisfies $X(p) = 0$, then $\Lambda(X)(p) = 0$.*

To see this, choose a chart $(\mathcal{U}, x)$ with $p \in \mathcal{U}$, and write $X = X^i\partial_i$ on $\mathcal{U}$, so the functions $X^i \in C^\infty(\mathcal{U})$ satisfy $X^1(p) = \ldots = X^n(p) = 0$. Using smooth cutoff functions, we can also choose global vector fields $e_1, \ldots, e_n \in \mathfrak{X}(M)$ and functions $f^1, \ldots, f^n \in C^\infty(M)$ such that

$$f^i = X^i \quad \text{and} \quad e_i = \partial_i \qquad \text{near } p, \text{ for all } i = 1, \ldots, n,$$

producing another vector field $Y := f^i e_i \in \mathfrak{X}(M)$ which matches X on some small neighborhood of p within $\mathcal{U}$. Claim 1 then implies $\Lambda(Y - X)(p) = \Lambda(Y)(p) - \Lambda(X)(p) = 0$. In light of C^∞-linearity and the condition $f^i(p) = X^i(p) = 0$ for $i = 1, \ldots, n$, we then have

$$\Lambda(X)(p) = \Lambda(Y)(p) = \Lambda(f^i e_i)(p) = f^i(p)\Lambda(e_i)(p) = 0.$$

□

From now on, we will say that a multilinear map on the spaces of vector fields and/or 1-forms **defines a tensor** whenever it is C^∞-linear in every argument, so that Proposition 8.2 or its obvious corollaries such as Exercise 8.3 apply. We can now carry out the "coordinate free" version of Exercise 8.1:

EXERCISE 8.4. Show that for any given 1-form $\lambda \in \Omega^1(M)$, the tensor of type $(0,2)$ that was defined via coordinates in Exercise 8.1 can also be defined via the bilinear map

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M) : (X,Y) \mapsto \mathcal{L}_X [\lambda(Y)] - \mathcal{L}_Y [\lambda(X)] - \lambda([X,Y]),$$

which is C^∞-linear in both arguments. (In this expression, we associate to each vector field $Z \in \mathfrak{X}(M)$ the smooth real-valued function $\lambda(Z) \in C^\infty(M)$ whose value at $p \in M$ is $\lambda(Z(p))$.)

EXERCISE 8.5. Suppose $J \in \Gamma(T^1_1 M)$ is a smooth almost complex structure, which we will regard as a smooth map $J : TM \to TM$ whose restriction to each tangent space T_pM is a linear map $J_p : T_pM \to T_pM$ with $J_p^2 = -\mathbb{1}$. The **Nijenhuis tensor**[32] is defined from J via the map

$$N : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \qquad N(X,Y) := [JX,JY] - J[JX,Y] - J[X,JY] - [X,Y].$$

(a) Use Exercise 8.3 to prove that this formula defines a tensor field of type $(1,2)$.
(b) Show that in local coordinates, the components of N and J are related by

$$N^i_{\ jk} = J^\ell_{\ j}\partial_\ell J^i_{\ k} - J^\ell_{\ k}\partial_\ell J^i_{\ j} + J^i_{\ \ell}\left(\partial_k J^\ell_{\ j} - \partial_j J^\ell_{\ k}\right).$$

(c) Show that N vanishes identically if $\dim M = 2$.
Hint: Notice that $N(X,Y)$ is antisymmetric in X and Y. What is $N(X,JX)$?
(d) An almost complex structure J is called *integrable* if near every point $p \in M$ there exists a chart $(\mathcal{U}, x)$ in which the components $J^i_{\ j}$ become the entries of the constant matrix

$$\mathbf{J}_0 := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \in \mathbb{R}^{2n\times 2n},$$

where each of the four blocks is an n-by-n matrix and $\dim M = 2n$. Show that if J is integrable, then $N \equiv 0$.
Advice: One can use the formula in part (b) for this, but an argument based directly on the definition of N via Lie brackets is also possible.

Remark: The matrix $\mathbf{J}_0$ represents the linear transformation $\mathbb{C}^n \to \mathbb{C}^n : \mathbf{z} \mapsto i\mathbf{z}$ if one identifies $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ via the correspondence $\mathbb{C}^n \ni \mathbf{x} + i\mathbf{y} \leftrightarrow (\mathbf{x},\mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$, thus an integrable almost complex structure makes M into a "complex manifold". By a deep theorem of Newlander and Nirenberg from 1957, the converse of part (d) is also true: if the Nijenhuis tensor vanishes, then J is integrable.

8.2. Differential forms and the exterior derivative. In Exercises 8.1 and 8.4, we saw that if we "antisymmetrize" the partial derivatives of the components of a 1-form, the result is a well-defined tensor field of type $(0,2)$. We shall now generalize this observation, and in the process, introduce an important special class of tensor fields that will play a major role when we discuss integration on manifolds.

A multilinear map $T : V \times \ldots \times V \to W$ is called **antisymmetric** (*antisymmetrisch*) or **skew-symmetric** (*schiefsymmetrisch*) or **alternating** if the value $T(v_1,\ldots,v_n)$ changes by a sign whenever any two of its arguments are interchanged. One can express this condition equivalently in terms of arbitrary permutations: let S_n denote the **symmetric group** on n elements, which consists of all bijections from the set $\{1,\ldots,n\}$ to itself, also known as **permutations** (*Permutationen*). There are exactly $n!$ elements in S_n, and the group is generated by the so-called *flips*,

[32]Approximate pronounciation: "NIGH-en-house", where "nigh" rhymes with English "sigh".

which satisfy $\sigma(i) = j$ and $\sigma(j) = i$ for two distinct elements $i, j \in \{1, \dots, n\}$ while leaving every other element fixed. Every permutation can therefore be expressed as a composition of flips, and while a given permutation will generally admit many distinct decompositions into varying numbers of flips, one can show that for any fixed $\sigma \in S_n$, the number of flips required is always either even or odd, i.e. a composition of evenly many flips cannot also be expressed as a composition of an odd number of flips, or vice versa. We call each permutation $\sigma \in S_n$ **even** (*gerade*) or **odd** (*ungerade*) accordingly, and define its **parity** by[33]

$$|\sigma| := \begin{cases} 0 & \text{if } \sigma \text{ is even,} \\ 1 & \text{if } \sigma \text{ is odd.} \end{cases}$$

In applications, the parity usually appears in the form $(-1)^{|\sigma|}$, thus one sometimes also refers to odd or even permutations as *negative* or *positive* respectively. With this notion in place, a multilinear map $T : \underbrace{V \times \ldots \times V}_{n} \to W$ is antisymmetric if and only if it satisfies

$$T(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = (-1)^{|\sigma|} T(v_1, \dots, v_n)$$

for all $v_1, \dots, v_n \in V$ and $\sigma \in S_n$. One can turn any multilinear map $T : V \times \ldots \times V \to W$ into one that is antisymmetric by defining

$$(\operatorname{Alt} T)(v_1, \dots, v_n) := \frac{1}{n!} \sum_{\sigma \in S_n} (-1)^{|\sigma|} T(v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

We observe that $\operatorname{Alt}(T) = T$ if and only if T is antisymmetric, thus Alt defines a linear projection map $\operatorname{Hom}(\bigotimes^n V, W) \to \operatorname{Hom}(\bigotimes^n V, W)$ onto the subspace of antisymmetric maps.

Definition 8.6. For any integer $k \geqslant 0$, an antisymmetric tensor field of type $(0, k)$ on M is called a **differential k-form** (or just **k-form** for short). The vector space of smooth k-forms on M is denoted by

$$\Omega^k(M) := \{\text{smooth } k\text{-forms on } M\}.$$

Note that antisymmetry is a vacuous condition in the cases $k = 0, 1$, which is why $\Omega^1(M) = \Gamma(T_1^0 M)$ and $\Omega^0(M) = \Gamma(T_0^0 M) = C^\infty(M)$. Given a chart $(\mathcal{U}, x)$, a k-form $\omega \in \Omega^k(M)$ can be written in local coordinates as

$$\omega = \omega_{i_1 \dots i_k} \, dx^{i_1} \otimes \ldots \otimes dx^{i_k} \qquad \text{on } \mathcal{U},$$

where antisymmetry means that the component functions $\omega_{i_1 \dots i_k} : \mathcal{U} \to \mathbb{R}$ change by a sign whenever two of the indices are interchanged. In this context, the following notational device is often useful. Suppose $T_{i_1 \dots i_k}$ is a collection of symbols associating to each k-tuple of integers $i_1, \dots, i_k \in \{1, \dots, n\}$ an element of some vector space, e.g. $C^\infty(\mathcal{U})$ in the example above. We can then **antisymmetrize** these symbols to define

$$T_{[i_1 \dots i_k]} := \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^{|\sigma|} T_{i_{\sigma(1)} \dots i_{\sigma(k)}},$$

so the symbols $T_{[i_1 \dots i_k]}$ are antisymmetric with respect to interchanging pairs of indices, and one has $T_{[i_1 \dots i_k]} = T_{i_1 \dots i_k}$ if and only if $T_{i_1 \dots i_k}$ already has this property. Note that in this definition, there is no need to assume that $T_{i_1 \dots i_k}$ are the components of a well-defined tensor, but usefully, it may nonetheless happen that $T_{[i_1 \dots i_k]}$ *does* define a tensor. We saw an example of this already

[33] One easy way to see that the parity is well defined is by associating to each permutation $\sigma \in S_n$ the unique linear map $\mathbf{A}_\sigma : \mathbb{R}^n \to \mathbb{R}^n$ that permutes the standard basis vectors by σ. The matrix of $\mathbf{A}_\sigma$ is obtained from the identity matrix by permuting its columns, and $\det \mathbf{A}_\sigma = (-1)^{|\sigma|}$.

in Exercise 8.1, where the tensor $S \in \Gamma(T_2^0 M)$ defined from any 1-form $\lambda \in \Omega^1(M)$ can now be abbreviated in local coordinates by

$$S_{ij} = 2\,\partial_{[i}\lambda_{j]}$$

PROPOSITION 8.7. *For every smooth differential form $\omega \in \Omega^k(M)$, $k \geqslant 0$, there exists a unique $(k+1)$-form $d\omega \in \Omega^{k+1}(M)$ determined by the formula*

$$d\omega(X_0, \ldots, X_k) = \sum_{i=0}^{k} (-1)^i \mathcal{L}_{X_i}\left[\omega(X_0, \ldots, \hat{X}_i, \ldots, X_k)\right] + \sum_{0 \leqslant i < j \leqslant k} (-1)^{i+j} \omega\big([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_k\big) \tag{8.4}$$

for $X_0, \ldots, X_k \in \mathfrak{X}(M)$, where the hats over certain terms in sequences like "$X_0, \ldots, \hat{X}_i, \ldots, X_k$" mean that those terms do not appear in the sequence but every other term does. For any chart $(\mathcal{U}, x)$, the components of $d\omega$ in local coordinates over $\mathcal{U} \subset M$ are given by

$$(d\omega)_{i_0 \ldots i_k} = (k+1)\partial_{[i_0}\omega_{i_1 \ldots i_k]}.$$

PROOF. We claim first that both terms on the right hand side of (8.4) are antisymmetric functions of the vector fields $X_0, \ldots, X_k$. In fact, the first term satisfies

$$\sum_{i=0}^{k} (-1)^i \mathcal{L}_{X_i}\left[\omega(X_0, \ldots, \hat{X}_i, \ldots, X_k)\right] = \frac{1}{k!} \sum_{\sigma \in S_{k+1}} (-1)^{|\sigma|} \mathcal{L}_{X_{\sigma(0)}}\left[\omega(X_{\sigma(1)}, \ldots, X_{\sigma(k)})\right], \tag{8.5}$$

where the right hand side is manifestly antisymmetric, and in this setting S_{k+1} means the group of permutations of the elements $\{0, \ldots, k\}$. This can be seen by considering separately for each $i = 0, \ldots, k$ the permutations σ with $\sigma(0) = i$, and then exploiting the antisymmetry of ω to place $X_{\sigma(1)}, \ldots, X_{\sigma(k)}$ in a canonical order. A similar approach shows that the second term is a constant multiple of the antisymmetric expression $\sum_{\sigma \in S_{k+1}} (-1)^{|\sigma|} \omega\big([X_{\sigma(0)}, X_{\sigma(1)}], X_{\sigma(2)}, \ldots, X_{\sigma(k)}\big)$.

We claim next that the right hand side of (8.4) is C^∞-linear in X_i for every $i = 0, \ldots, k$. By antisymmetry, it suffices to prove this for $i = 0$, and the proof is then a straightforward computation based on Exercise 6.4. We can now conclude from Proposition 8.2 that $d\omega$ is a well-defined $(k+1)$-form. Finally, the coordinate formula for $d\omega$ follows from (8.5) since $[\partial_i, \partial_j] \equiv 0$ for all i, j. ☐

DEFINITION 8.8. For a smooth k-form on ω, the $(k+1)$-form $d\omega$ defined in Proposition 8.7 is called the **exterior derivative** (*äußere Ableitung*) of ω.

EXAMPLE 8.9. For a 0-form $f \in C^\infty(M) = \Omega^0(M)$, the definition above makes $df \in \Omega^1(M)$ the usual differential of f.

For $k > 0$, the exterior derivative $d\omega$ of $\omega \in \Omega^k(M)$ does not contain *all* information about the first derivative of ω at each point, e.g. in local coordinates, the individual partial derivatives $\partial_j \omega_{i_1 \ldots i_k}$ cannot be deduced from $(d\omega)_{i_0 \ldots i_k}$, nor can ω be recovered from $d\omega$ up to addition of a constant. We will see more comprehensive (though non-canonical) ways of defining derivatives of ω when we discuss connections. The exterior derivative will be essential, however, due to the role it plays in Stokes' theorem, the n-dimensional generalization of the fundamental theorem of calculus.

8.3. Pullbacks and pushforwards. For a diffeomorphism $\psi : M \to N$, pushforwards and pullbacks of tensor fields can be defined in much the same way as for functions and vector fields in §5.2. Recalling the notation

$$\psi_* := T\psi : TM \to TN, \qquad \psi^* := (T\psi)^{-1} : TN \to TM,$$

we can dualize to define

$$\psi^* : T^*N \to T^*M, \qquad \psi_* : T^*M \to T^*N$$

by

$$(\psi^*\lambda)(X) := \lambda(\psi_* X), \qquad (\psi_*\lambda)(X) := \lambda(\psi^* X).$$

Every $S \in \Gamma(T^k_\ell M)$ with $k > 0$ or $\ell > 0$ then has a **pushforard** $\psi_* S \in \Gamma(T^k_\ell N)$ defined by

$$(\psi_* S)(\lambda^1, \dots, \lambda^k, X_1, \dots, X_\ell) := S(\psi^*\lambda^1, \dots, \psi^*\lambda^k, \psi^* X_1, \dots, \psi^* X_\ell),$$

and similarly, $S \in \Gamma(T^k_\ell N)$ has a **pullback** $\psi^* S \in \Gamma(T^k_\ell M)$ defined by

$$(\psi^* S)(\lambda^1, \dots, \lambda^k, X_1, \dots, X_\ell) := S(\psi_*\lambda^1, \dots \psi_*\lambda^k, \psi_* X_1, \dots, \psi_* X_\ell).$$

The reader should take a moment to check that under the canonical identification $\mathfrak{X}(M) = \Gamma(T^1_0 M)$, this definition of the pushforward and pullback for tensor fields of type $(1, 0)$ matches what we defined in §5.2 for vector fields. The maps

$$\psi_* : \Gamma(T^k_\ell M) \to \Gamma(T^k_\ell N), \qquad \psi^* : \Gamma(T^k_\ell N) \to \Gamma(T^k_\ell M)$$

are vector space isomorphisms, and are inverse to each other. It is straightforward to show that if $\varphi : N \to Q$ is another diffeomorphism, the composition $\varphi \circ \psi : M \to Q$ satisfies

$$(\varphi \circ \psi)_* = \varphi_* \psi_*, \qquad (\varphi \circ \psi)^* = \psi^* \varphi^*. \tag{8.6}$$

Notice that the pushforward $\psi_* X = T\psi(X) \in TN$ of a tangent vector $X \in TM$ is defined without reference to the inverse ψ^{-1}, and can therefore also be defined when $\psi : M \to N$ is any smooth map, not necessarily a diffeomorphism. The same thus holds for the *pullback* of a fully covariant tensor field $S \in \Gamma(T^0_k N)$: the definition of $\psi^* S \in \Gamma(T^0_k M)$ as

$$\psi^* S(X_1, \dots, X_\ell) = S(\psi_* X_1, \dots, \psi_* X_\ell) = S(T\psi(X_1), \dots, T\psi(X_\ell))$$

makes sense for any smooth map $\psi : M \to N$, though the resulting linear map $\psi^* : \Gamma(T^0_k N) \to \Gamma(T^0_k M)$ need not be invertible if ψ is not a diffeomorphism. This applies in particular for differential forms: they can always be pulled back via smooth maps.

Exercise 8.10. Assume $\psi : M \to N$ is a smooth map and $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ are charts on M and N respectively such that $\mathcal{U} \cap \psi^{-1}(\mathcal{V}) \neq \varnothing$. Abbreviating $\psi^i := y^i \circ \psi : \psi^{-1}(\mathcal{V}) \to \mathbb{R}$ for the component functions of ψ written in coordinates, show that the components of a k-form $\omega \in \Omega^k(N)$ in the coordinates $y^1, \dots, y^n$ are related to those of its pullback $\psi^*\omega \in \Omega^k(M)$ in coordinates $x^1, \dots, x^m$ by

$$(\psi^*\omega)_{i_1 \dots i_k} = \frac{\partial \psi^{j_1}}{\partial x^{i_1}} \dots \frac{\partial \psi^{j_k}}{\partial x^{i_k}} (\omega_{j_1 \dots j_k} \circ \psi) \qquad \text{on } \mathcal{U} \cap \psi^{-1}(\mathcal{V}).$$

8.4. The Lie derivative of a tensor field. As with vector fields in §6.4, there is a natural way to differentiate any tensor field $S \in \Gamma(T^k_\ell M)$ with respect to a vector field $X \in \mathfrak{X}(M)$, giving the most general version of the **Lie derivative**

$$\mathcal{L}_X S := \left. \frac{d}{dt} (\varphi^t_X)^* X \right|_{t=0} \in \Gamma(T^k_\ell M).$$

This is well defined even if none of the flow maps φ^t_X are globally defined on M for $t \neq 0$, since for any point $p \in M$, φ^t_X is at least defined on a neighborhood of p for every t close enough to 0.

As with the Lie derivative of vector fields, one should keep in mind that for each $p \in M$, $(\mathcal{L}_X S)_p$ depends on more than just S and the value of X at p, due to the fact that pulling back via the flow requires differentiating it, and this derivative will also depend on the derivatives of X at p. The only exception is the case $k = \ell = 0$, in which S is just a function $f : M \to \mathbb{R}$ and $\mathcal{L}_X f = df(X)$ as before.

The Lie derivative has important applications to questions of *invariance*, e.g. if $\dim M = n$, we will see that one can use a differential form $\omega \in \Omega^n(M)$ to define a notion of *volume* for regions in M, and the condition $\mathcal{L}_X \omega \equiv 0$ will then characterize vector fields whose flows are volume preserving. We will need to develop the technology somewhat further before we can do nontrivial things with this, as it is typically quite difficult to compute $\mathcal{L}_X S$ directly from the definition, due to the fact that the flow of a vector field is typically not easy to write down. Let us mention however that there is a very user-friendly formula for the Lie derivative of a differential form:

THEOREM 8.11 (Cartan's formula). *For any $\omega \in \Omega^k(M)$ and $X \in \mathfrak{X}(M)$,*

$$\mathcal{L}_X \omega = d(\iota_X \omega) + \iota_X(d\omega),$$

where the ***interior product*** *$\iota_X \alpha \in \Omega^{q-1}(M)$ of a differential form $\alpha \in \Omega^q(M)$ with a vector field $X \in \mathfrak{X}(M)$ is defined by*

$$(\iota_X \alpha)(Y_1, \ldots, Y_{q-1}) := \alpha(X, Y_1, \ldots, Y_{q-1}).$$

We will prove this in Lecture 11, after we have discussed the algebra of differential forms in more detail.

9. The algebra of differential forms

Our goal for the next two lectures is to make sense of symbols like $\int_M f$ when M is a manifold. The naive hope would be that one could associate a real number $\int_M f \in \mathbb{R}$ to every (let's say continuous and compactly supported) function $f : M \to \mathbb{R}$, one that weights the values of f in proportion to the amount of volume covered. We will see that this notion does not make sense in general for real-valued *functions*, but if $\dim M = n$, it does make sense when f is replaced by a differential n-form.

9.1. Measure and volume on manifolds. The basic problem with defining $\int_M f$ for a function $f : M \to \mathbb{R}$ is that we have not specified any measure on M with which to define what "volume" means. Certain special classes of manifolds admit canonical measures, e.g. if M is a k-dimensional submanifold of $\mathbb{R}^n$, then one can derive a notion of "k-dimensional volume" on subsets of M from the Euclidean geometry of $\mathbb{R}^n$. But this measure on M will depend on the precise embedding $M \hookrightarrow \mathbb{R}^n$, e.g. the volume of any given region in M will change by a factor of L^k if we modify the embedding by multiplication with a scalar $L > 0$. And in any case, not all manifolds are presented as submanifolds of Euclidean space.

Another idea would be to use local coordinates, meaning that for any chart $(x, \mathcal{U})$ on M, the measure of a subset $\mathcal{O} \subset \mathcal{U}$ could be defined as the Lebesgue measure of $x(\mathcal{O}) \subset \mathbb{R}^n$. This definition, however, clealy depends on the choice of chart: according to the change of variables formula, the Lebesgue measure of $y(\mathcal{O}) \subset \mathbb{R}^n$ for another chart $(\mathcal{V}, y)$ with $\mathcal{O} \subset \mathcal{V}$ will be the Lebesgue integral of $|\det D(y \circ x^{-1})|$ over $x(\mathcal{O})$, and this integral is not typically the same as the measure of $x(\mathcal{O})$.

Let us drop the question of whether M carries a canonical measure (usually it doesn't), and ask instead how one might go about *choosing* a measure on M, i.e. what kinds of properties should a notion of n-dimensional volume on M have? Heuristically, one useful way to approach this question is by thinking of the tangent space T_pM at a point $p \in M$ is an "approximation" of a neighborhood of p in M, so if we can define volumes of regions in that neighborhood, we should also be able to define volumes of regions in the vector space T_pM. How does one define volume in an n-dimensional vector space? For example, given vectors $X_1, \dots, X_n \in T_pM$, consider the so-called **parallelepiped** spanned by $X_1, \dots, X_n$, meaning the set

$$P(X_1, \dots, X_n) := \left\{ t^i X_i \in T_pM \;\middle|\; t^1, \dots, t^n \in [0,1] \right\} \subset T_pM,$$

where as usual there is an implied summation in the expression $t^i X_i$. Suppose $\mu : T_pM \times \dots \times T_pM \to [0, \infty)$ is a function that associates to each n-tuple $(X_1, \dots, X_n)$ the n-dimensional volume of $P(X_1, \dots, X_n)$. What kind of function is μ? Basic geometric considerations dictate the following:

(1) If one of the vectors X_i is multiplied by a nonnegative constant, the volume scales by the same constant, i.e.

$$\mu(X_1, \dots, cX_i, \dots, X_n) = c\mu(X_1, \dots, X_i, \dots, X_n)$$

for $c \geqslant 0$.

(2) The volume is additive[34] with respect to each variable, i.e.

$$\mu(X_1, \dots, X_i + X_i', \dots, X_n) = \mu(X_1, \dots, X_i, \dots, X_n) + \mu(X_1, \dots, X_i', \dots, X_n).$$

An elementary geometric justification of this relation in the case $n = 2$ is shown in Figure 7. Using the letters A through E to denote the areas of the various regions in this picture, one has $\mu(X_1, X_2) = A + B$, $\mu(X_1', X_2) = C + D$, and $\mu(X_1 + X_1', X_2) = A + C + E = A + C + B + D = \mu(X_1, X_2) + \mu(X_1', X_2)$.

(3) If any two of the vectors $X_1, \dots, X_n$ match, then $P(X_1, \dots, X_n)$ is contained in an $(n-1)$-dimensional subspace and thus has zero n-dimensional volume, so

$$\mu(X_1, \dots, X_n) = 0 \quad \text{whenever} \quad X_i = X_j \text{ for some } i \neq j.$$

The first two properties suggest multilinearity, though μ itself cannot be multilinear since it only takes nonnegative values, and the scalar multiplication property only involves nonnegative scalars. On the other hand, a good way to find functions μ that satisfy these two properties is by choosing an actual multilinear function $\omega : T_pM \times \dots \times T_pM \to \mathbb{R}$ and setting

$$\mu(X_1, \dots, X_n) := |\omega(X_1, \dots, X_n)|.$$

The third property now imposes a serious restriction on ω:

PROPOSITION 9.1. *If V is a vector space and $\omega : V \times \dots \times V \to \mathbb{R}$ is an n-fold multilinear function that vanishes whenever two of its arguments are identical, then ω is alternating.*

PROOF. In the case $n = 2$, it suffices to choose any $v, w \in V$ and use multilinearity to observe

$$0 = \omega(v + w, v + w) = \omega(v, v) + \omega(w, w) + \omega(v, w) + \omega(w, v) = \omega(v, w) + \omega(w, v).$$

The general case works similarly. □

The upshot of this discussion is that a reasonable notion of volume for paralelepipeds in a tangent space T_pM can be defined by choosing an alternating n-fold multilinear form ω on T_pM and taking its absolute value. If the gaps in the discussion leading to this conclusion made you uncomfortable, one could alternatively derive it from a basic result in measure theory: every translation-invariant measure on $\mathbb{R}^n$ is a scalar $c \geqslant 0$ multiplied by the Lebesgue measure (see e.g. [**Sal16**, Chapter 2]). Moreover, the Lebesgue measure of the parallelepiped spanned by n vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in $\mathbb{R}^n$ is given by $\left|\det \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}\right|$. As you learned in linear algebra, the

[34] Strictly speaking, some extra condition on the vectors $X_1, \dots, X_n$ is needed in order for the additivity property to hold, as not all possible configurations (even in the case $n = 2$) can be described by something like Figure 7. Since this is only meant to be a heuristic discussion, let's not worry about this for now.

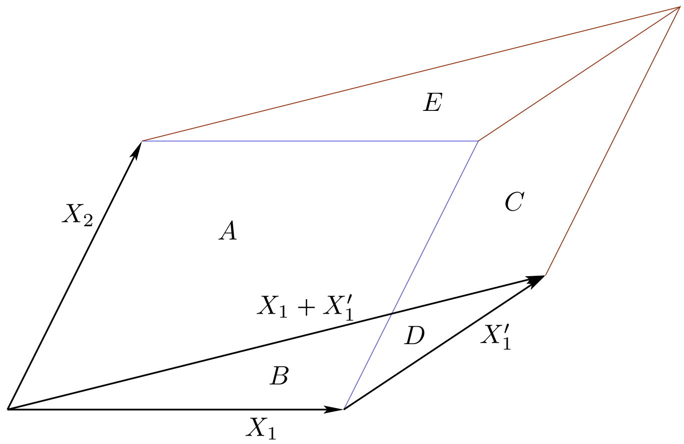


FIGURE 7. A geometric "proof" that volumes of parallelepipeds are determined by *multilinear* functions of their spanning vectors.

determinant of a matrix is an alternating multilinear function of its columns, thus we can now write $\mu = |\omega|$ where $\omega(\mathbf{v}_1, \dots, \mathbf{v}_n) := c \det \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$ defines an alternating multilinear form.

Since everything in this course is smooth, it will also make sense to assume that for reasonable notions of volume on regions in M, the associated notions of volume on the tangent spaces T_pM depend smoothly on the point p. We can now say precisely what kind of geometric object defines a smoothly varying notion of volume on tangent spaces: it is a smooth n-form $\omega \in \Omega^n(M)$.

9.2. Exterior algebra. The previous section provided some motivation to believe that differential forms are the right objects with which to define integration on manifolds. Before we can fully unpack this idea, we need to develop the algebra of differential forms a bit further.

The tasks of this section are fundamentally algebraic, so there will be no manifolds, only an n-dimensional vector space V with basis $e_1, \dots, e_n \in V$. Let $e^1_*, \dots, e^n_* \in V^*$ denote the corresponding **dual basis**, determined by the condition

$$e^i_*(e_j) = \delta^i_j.$$

Recall from §7.2 that V^k_ℓ denotes the space of multilinear functions $V^* \times \dots \times V^* \times V \times \dots \times V \to \mathbb{R}$ that take k dual vectors in V^* and ℓ vectors in V as arguments; in particular, $V^0_1 = V^*$ and V^1_0 is the "double dual" $(V^*)^*$ of V, which is canonically isomorphic to V itself. The tensor product $\otimes : V^k_\ell \times V^r_s \to V^{k+r}_{\ell+s}$ can be defined in the same way as for tensor fields, and it is associative, so in particular, the tensor product of k dual vectors $\alpha^1, \dots, \alpha^k$ is a k-fold multilinear map $\alpha^1 \otimes \dots \otimes \alpha^k : V \times \dots \times V \to \mathbb{R}$ defined by

$$(\alpha^1 \otimes \dots \otimes \alpha^k)(v_1, \dots, v_k) = \alpha^1(v_1) \cdot \ldots \cdot \alpha^k(v_k).$$

The vector space of real-valued alternating k-fold multilinear maps on V is denoted by

$$\Lambda^k V^* := \left\{ \omega \in V^0_k \;\middle|\; \omega(\dots, v, \dots, w, \dots) = -\omega(\dots, w, \dots, v, \dots) \text{ for all } v, w \in V \right\},$$

and we often refer to its elements as **alternating k-forms on V**. The antisymmetry condition is vacuous for $k \leqslant 1$, thus $\Lambda^0 V^* = \mathbb{R}$ and $\Lambda^1 V^* = V^*$. Using multilinearity as in Proposition 7.20, any $\omega \in \Lambda^k V^*$ for $k \geqslant 1$ can be written in terms of the basis $e^1_*, \dots, e^n_* \in V^*$ as

$$\omega = \omega_{i_1 \dots i_k} \, e^{i_1}_* \otimes \dots \otimes e^{i_k}_*,$$

with unique coefficients

$$\omega_{i_1\dots i_k} := \omega(e_{i_1},\dots,e_{i_k}) \in \mathbb{R}. \tag{9.1}$$

These coefficients are not all independent of each other: the antisymmetry of ω dictates that they satisfy

$$\omega_{i_1\dots j\dots\ell\dots i_k} = -\omega_{i_1\dots\ell\dots j\dots i_k},$$

i.e. there is a sign change whenever two distinct indices are interchanged, and $\omega_{i_1\dots i_k}$ can only be nontrivial when all of its indices $i_1,\dots,i_k \in \{1,\dots,n\}$ have distinct values. It follows that $\omega_{i_1\dots i_k}$ must always vanish if $k > n$, and otherwise, the number of distinct components that can be specified independently before the rest are determined is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, hence

$$\dim \Lambda^k V^* = \begin{cases} \binom{n}{k} = \frac{n!}{k!(n-k)!} & \text{for } k \leqslant n, \\ 0 & \text{for } k > n. \end{cases}$$

Observe that while the case $k = 0$ was excluded from the discussion above, the formula $\dim \mathbb{R} = \dim \Lambda^0 V^* = \binom{n}{0} = 1$ is also correct in that case. The most interesting case is $k = n$: the elements of $\Lambda^n V^*$ are sometimes called **top-dimensional** forms, since n is the largest value of k for which $\Lambda^k V^*$ is a nontrivial space. The space is 1-dimensional in this case, due to the fact that all nontrivial components of $\omega \in \Lambda^n V^*$ are obtained by permuting the indices of $\omega_{1\dots n}$. This elementary observation has nontrivial consequences that will be concretely useful to us, such as:

PROPOSITION 9.2. *For any basis $v_1,\dots,v_n \in V$ of a vector space V, every $\omega \in \Lambda^n V^*$ is uniquely determined by the number $\omega(v_1,\dots,v_n) \in \mathbb{R}$; in particular, this number vanishes if and only if $\omega = 0$.* □

EXAMPLE 9.3. The **determinant** $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ can be characterized by the property that $\mathbb{R}^n \times \dots \times \mathbb{R}^n \to \mathbb{R} : (\mathbf{v}_1,\dots,\mathbf{v}_1) \mapsto \det\begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{pmatrix}$ is the unique element of $\Lambda^n(\mathbb{R}^n)^*$ satisfying $\det\begin{pmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_n\end{pmatrix} = 1$ for the standard basis $\mathbf{e}_1,\dots,\mathbf{e}_n \in \mathbb{R}^n$. Using the dual basis $\mathbf{e}_*^1,\dots,\mathbf{e}_*^n \in (\mathbb{R}^n)^*$ to the standard basis, one can write down a concrete element of $\Lambda^n(\mathbb{R}^n)^*$ with this property in the form

$$\sum_{\sigma\in S_n} (-1)^{|\sigma|} \mathbf{e}_*^{\sigma(1)} \otimes \dots \otimes \mathbf{e}_*^{\sigma(n)} \in \Lambda^n(\mathbb{R}^n)^*.$$

Plugging in the columns of a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ with entries $A^i{}_j$, an explicit formula for the determinant is thus given by

$$\det(\mathbf{A}) = \sum_{\sigma\in S_n} (-1)^{|\sigma|} A^{\sigma(1)}{}_1 \cdot \dots \cdot A^{\sigma(n)}{}_n. \tag{9.2}$$

Proposition 9.2 now implies that every $\omega \in \Lambda^n(\mathbb{R}^n)^*$ can be written as

$$\omega(\mathbf{v}_1,\dots,\mathbf{v}_n) = c \cdot \det\begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{pmatrix},$$

with a constant given by $c := \omega(\mathbf{e}_1,\dots,\mathbf{e}_n) \in \mathbb{R}$.

For $k \geqslant 1$, a natural linear projection $\mathrm{Alt} : V_k^0 \to V_k^0$ onto the subspace $\Lambda^k V^* \subset V_k^0$ is defined by

$$\mathrm{Alt}(\omega)(v_1,\dots,v_k) := \frac{1}{k!}\sum_{\sigma\in S_k} (-1)^{|\sigma|}\omega(v_{\sigma(1)},\dots,v_{\sigma(k)}).$$

Indeed, one readily checks that $\mathrm{Alt}(\omega)$ is alternating for every $\omega \in V_k^0$, and ω itself is alternating if and only if $\mathrm{Alt}(\omega) = \omega$. If we write $\omega = \omega_{i_1\dots i_k}\, e_*^{i_1} \otimes \dots \otimes e_*^{i_k}$ for a general $\omega \in V_k^0$, applying

Alt changes the components via the antisymmetrization operation introduced in §8.2, which can be written succinctly as

$$\mathrm{Alt}(\omega)_{i_1\dots i_k} = \omega_{[i_1\dots i_k]}.$$

Note that for $k = 1$, Alt is simply the identity map $V^* \to V^*$. It will be a useful convention to extend this definition to $k = 0$ so that Alt is also the identity map on $V^0_0 = \mathbb{R}$.

We would now like to define a product operation on alternating forms that has geometric meaning. Let us regard each of the chosen basis 1-forms $e^i_* \in \Lambda^1 V^*$ as defining a notion of *length* (also known as "1-dimensional volume") for vectors in the 1-dimensional subspace $V_i := \mathbb{R}e_i \subset V$, so by this definition, the basis vectors $e_i \in V_i$ have unit length. The fact that each e^i_* vanishes on all the other subspaces $V_j \subset V$ for $j \neq i$ can be interpreted moreover as an "orthogonality" condition, so that we regard all the subspaces $V_1, \dots, V_n \subset V$ as orthogonal to each other. Geometrically, the paralelepiped in V spanned by $e_1, \dots, e_n$ should then have volume 1, and we would like to define the product n-form $e^1_* \wedge \dots \wedge e^n_* \in \Lambda^n V^*$ to reproduce this notion of volume, i.e. it should satisfy

$$(e^1_* \wedge \dots \wedge e^n_*)(e_1, \dots, e_n) = 1.$$

Since $\dim \Lambda^n V^* = 1$, there is exactly one element of $\Lambda^n V^*$ that satisfies this condition, and it is given by

$$e^1_* \wedge \dots \wedge e^n_* = n!\,\mathrm{Alt}(e^1_* \otimes \dots \otimes e^n_*) = \sum_{\sigma \in S_n} (-1)^{|\sigma|} e^{\sigma(1)}_* \otimes \dots \otimes e^{\sigma(n)}_*.$$

We take this observation as motivation for the general definition of the **wedge product**, which is contained in the theorem below. To state it properly, we define the vector space

$$\Lambda^* V^* := \bigoplus_{k=0}^{\infty} \Lambda^k V^*,$$

which is finite dimensional since $\Lambda^k V^* = \{0\}$ for $k > n$, hence $\Lambda^* V^*$ is equivalent to the finite product $\Lambda^0 V^* \times \dots \times \Lambda^n V^*$. We can regard each of the spaces $\Lambda^k V^*$ as subspaces of $\Lambda^* V^*$ in the obvious way. A nontrivial element $\alpha \in \Lambda^* V^*$ is said to be **homogeneous of degree** k if it belongs to the subspace $\Lambda^k V^* \subset \Lambda^* V^*$, in which case we also sometimes write its degree as

$$\deg(\alpha) = |\alpha| := k \quad \text{for} \quad \alpha \in \Lambda^k V^*.$$

One should keep in mind that not all elements of $\Lambda^* V^*$ are homogeneous, but this is of little importance in practice because every nontrivial element is a sum of a unique finite set of homogeneous elements of various degrees.

THEOREM 9.4. *There exists a unique bilinear map* $\Lambda^* V^* \times \Lambda^* V^* \to \Lambda^* V^* : (\alpha, \beta) \mapsto \alpha \wedge \beta$ *that satisfies*

$$c \wedge \alpha = \alpha \wedge c := c\alpha \qquad \textit{for all } \alpha \in \Lambda^* V^* \textit{ and } c \in \Lambda^0 V^* = \mathbb{R},$$

the associativity property

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) \qquad \textit{for all } \alpha, \beta, \gamma \in \Lambda^* V^*,$$

and

$$\alpha^1 \wedge \dots \wedge \alpha^k = \sum_{\sigma \in S_k} (-1)^{|\sigma|} \alpha^{\sigma(1)} \otimes \dots \otimes \alpha^{\sigma(k)} \qquad \textit{for all } k \in \mathbb{N},\ \alpha^1, \dots, \alpha^k \in \Lambda^1 V^*, \tag{9.3}$$

where the k-fold product on the left hand side is defined by arbitrarily inserting parentheses to produce a sequence of binary operations. Moreover, the following conditions are satisfied:

(1) For any integers $k, \ell \geqslant 0$ and $\alpha \in \Lambda^k V^*$, $\beta \in \Lambda^\ell V^*$,

$$\alpha \wedge \beta = \frac{(k+\ell)!}{k!\ell!} \mathrm{Alt}(\alpha \otimes \beta) \in \Lambda^{k+\ell} V^*. \tag{9.4}$$

(2) *The wedge product is* graded commutative, *i.e. for homogeneous elements* $\alpha, \beta \in \Lambda^* V^*$,

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

Before proving the theorem, we make the useful observation that if one defines k-fold wedge products of 1-forms via the right hand side of (9.3), then they can be used to turn any basis of V^* into a basis of $\Lambda^k V^*$:

PROPOSITION 9.5. *Given the basis* $e_1, \ldots, e_n \in V$ *and its dual basis* $e^1_*, \ldots, e^n_* \in V^*$, *every* $\omega \in \Lambda^k V^*$ *can be written as*

$$\omega = \sum_{i_1 < \ldots < i_k} \omega_{i_1 \ldots i_k} \, e^{i_1}_* \wedge \ldots \wedge e^{i_k}_* \tag{9.5}$$

for unique coefficients $\omega_{i_1 \ldots i_k} \in \mathbb{R}$, *which are given by*[35]

$$\omega_{i_1 \ldots i_k} = \omega(e_{i_1}, \ldots, e_{i_k}) \in \mathbb{R}.$$

PROOF. One uses the formula (9.3) to show that both sides of (9.5) match when evaluated on any tuple of basis vectors $(e_{i_1}, \ldots, e_{i_k})$ with $i_1 < \ldots < i_k$, and by antisymmetry, it follows that they also match when evaluated on *any* tuple of basis vectors. Multilinearity then implies that they match when evaluated on arbitrary k-tuples of vectors. □

REMARK 9.6. Proposition 9.5 is one of the few places where we are *not* using the Einstein summation convention. The reason is that the summation here does not cover all choices of tuples $i_1, \ldots, i_k \in \{1, \ldots, n\}$, as the summation convention would dictate, but rather only those for which the $i_1, \ldots, i_k$ are in strictly increasing order. Including all permutations of such tuples would produce extra terms that (due to the antisymmetry of both $\omega_{i_1 \ldots i_k}$ and $e^{i_1}_* \wedge \ldots \wedge e^{i_k}_*$) match the terms already present in the sum, i.e. exactly $k!$ copies of each term, plus some trivial terms for tuples in which some of the indices $i_1, \ldots, i_k$ match. This overcounting results in the formula

$$\omega = \frac{1}{k!} \omega_{i_1 \ldots i_k} \, e^{i_1}_* \wedge \ldots \wedge e^{i_k}_*,$$

in which the coefficients are defined the same as before but the summation convention *is* in effect.

EXAMPLE 9.7. The following case of (9.3) is worth drawing special attention to: for two 1-forms $\alpha, \beta \in \Lambda^1 V^*$, $\alpha \wedge \beta \in \Lambda^2 V^*$ is given by $\alpha \wedge \beta = \alpha \otimes \beta - \beta \otimes \alpha$, thus

$$(\alpha \wedge \beta)(v, w) = \alpha(v)\beta(w) - \beta(v)\alpha(w).$$

One sees easily from this formula that the wedge product of 1-forms is *anticommutative*, i.e. it satisfies $\alpha \wedge \beta = -\beta \wedge \alpha$, and in particular, $\alpha \wedge \alpha = 0$.

PROOF OF THEOREM 9.4. By Proposition 9.5, every $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^\ell V^*$ for $k, \ell \geqslant 1$ can be expressed as sums of wedge products of the basis 1-forms $e^1_*, \ldots, e^n_* \in V^*$ as determined by (9.3), so bilinearity and associativity together with (9.3) then uniquely determine $\alpha \wedge \beta \in \Lambda^{k+\ell} V^*$. The only problem with taking the resulting formula as a general *definition* of $\alpha \wedge \beta$ is that it may a priori depend on the choice of the basis $e^1_*, \ldots, e^n_*$. In order to dismiss this concern, we will show that this definition of $\alpha \wedge \beta$ also satisfies the formula (9.4), and observe that the right hand side of this expression is clearly independent of choices. By bilinearity and Proposition 9.5, it suffices to check that this is true when α and β are themselves products of the form

$$\alpha = e^{i_1}_* \wedge \ldots \wedge e^{i_k}_*, \qquad \beta = e^{j_1}_* \wedge \ldots \wedge e^{j_\ell}_*$$

[35] Notice that the coefficients in Proposition 9.5 are the same ones that appeared in (9.1).

for some choice of $i_1, \ldots, i_k, j_1, \ldots, j_\ell \in \{1, \ldots, n\}$, and to show this, it is enough to evaluate both $\alpha \wedge \beta$ (as defined via (9.3)) and the right hand side of (9.4) on the ordered tuple of basis vectors

$$e_{a_1}, \ldots, e_{a_k}, e_{b_1}, \ldots, e_{b_\ell} \in V$$

for an arbitrary choice of $a_1, \ldots, a_k, b_1, \ldots, b_\ell \in \{1, \ldots, n\}$. By antisymmetry, both clearly vanish unless the integers $a_1, \ldots, a_k, b_1, \ldots, b_\ell$ are all distinct, so let us assume this. Both will also vanish if any of those numbers are not contained in the set $\{i_1, \ldots, i_k, j_1, \ldots, j_\ell\}$, so assume this as well from now on, which implies that the numbers $i_1, \ldots, i_k, j_1, \ldots, j_\ell$ must also be all distinct, and thus

$$\{a_1, \ldots, a_k, b_1, \ldots, b_\ell\} = \{i_1, \ldots, i_k, j_1, \ldots, j_\ell\}.$$

Using antisymmetry, we can now apply a permutation and assume without loss of generality that the two ordered tuples are exactly the same, i.e. $a_m = i_m$ and $b_m = j_m$ for all m, so we need only evaluate both $\alpha \wedge \beta$ and $\frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(\alpha \otimes \beta)$ on the ordered tuple

$$(v_1, \ldots, v_{k+\ell}) := e_{i_1}, \ldots, e_{i_k}, e_{j_1}, \ldots, e_{j_\ell}.$$

The result for $\alpha \wedge \beta$ is immediate from (9.3): only the trivial permutation produces a nontrivial term, and the answer is 1. Now consider

$$\frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(\alpha \otimes \beta)(v_1, \ldots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (-1)^{|\sigma|} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \cdot \beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}).$$

Since the sets $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_\ell\}$ are disjoint, the only permutations that contribute nontrivially to the right hand side of this expression are those which preserve the subsets $\{1, \ldots, k\}$ and $\{k+1, \ldots, k+\ell\}$, and the sign of such a permutation is the product of the signs of the permutations of these two subsets, so the sum can be rewritten as

$$\frac{1}{k!\ell!} \sum_{(\sigma_1,\sigma_2) \in S_k \times S_\ell} (-1)^{|\sigma_1|} \alpha(e_{i_{\sigma_1(1)}}, \ldots, e_{i_{\sigma_1(k)}}) \cdot (-1)^{|\sigma_2|} \beta(e_{j_{\sigma_2(1)}}, \ldots, e_{j_{\sigma_2(\ell)}}).$$

Finally, observe that since α and β are both antisymmetric, every term in this last sum is identical, and there are exactly $k!\ell!$ of them, so we can restrict to the trivial permutation and simplify the expression to

$$\alpha(e_{i_1}, \ldots, e_{i_k}) \cdot \beta(e_{j_1}, \ldots, e_{j_\ell}) = 1,$$

since both terms in the product equal 1 by (9.3). This establishes the existence of the associative product $\wedge : \Lambda^* V^* \times \Lambda^* V^* \to \Lambda^* V^*$ and the formula (9.4). One still has to show that it also satisfies (9.3), i.e. not just for the basis 1-forms e^i_* but for arbitrary tuples of 1-forms $\alpha^1, \ldots, \alpha^k \in \Lambda^1 V^*$. This can be derived from (9.4) by induction on k and a bit of combinatorics; we leave the details as an exercise.

To prove graded commutativity, it suffices again to consider the case where α and β are both products of 1-forms, and the relation then follows from the case $k = \ell = 1$ which was observed in Example 9.7. The key observation is that the number of flips required for permuting $i_1, \ldots, i_k, j_1, \ldots, j_\ell$ to $j_1, \ldots, j_\ell, i_1, \ldots, i_k$ is $k\ell$. □

The wedge product turns the vector space $\Lambda^* V^*$ into an algebra; it is called the **exterior algebra** (*äußere Algebra*) over V^*.[36]

[36] You may at this point be wondering what the "exterior algebra over V", presumably denoted by $\Lambda^* V$, might be. Since V is finite dimensional, the cheap way to define it is by identifying V with the dual space of V^*, so that homogeneous elements of $\Lambda^* V$ are antisymmetric multilinear maps $V^* \times \ldots \times V^* \to \mathbb{R}$. That is a correct definition, but not the most elegant formulation possible, and it also does not generalize to the case where V is infinite-dimensional since it may then fail to be isomorphic to its double dual. One can define $\Lambda^* V$ in terms of the abstract tensor product of vector spaces, and the details can be found in many standard algebra textbooks, but we will not need them here.

EXERCISE 9.8. Prove that a set of dual vectors $\alpha^1, \ldots, \alpha^k \in V^*$ is linearly independent if and only if its wedge product $\alpha^1 \wedge \ldots \wedge \alpha^k \in \Lambda^k V^*$ is nonzero.
Hint: Consider products of the form $\left(\sum_{i=1}^k c_i \alpha^i\right) \wedge \alpha^2 \wedge \ldots \wedge \alpha^k$.

EXERCISE 9.9. Show that if $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^\ell V^*$ are written in terms of the basis $e_*^1, \ldots, e_*^n \in V^*$ as $\alpha = \alpha_{i_1 \ldots i_k} e_*^{i_1} \otimes \ldots \otimes e_*^{i_k}$ and $\beta = \beta_{i_1 \ldots i_\ell} e_*^{i_1} \otimes \ldots \otimes e_*^{i_\ell}$, then $\alpha \wedge \beta = (\alpha \wedge \beta)_{i_1 \ldots i_{k+\ell}} e_*^{i_1} \otimes \ldots \otimes e_*^{i_{k+\ell}}$ where

$$(\alpha \wedge \beta)_{i_1 \ldots i_{k+\ell}} = \frac{(k+\ell)!}{k!\ell!} \alpha_{[i_1 \ldots i_k} \beta_{i_{k+1} \ldots i_{k+\ell}]}.$$

The following formula for top-dimensional forms will have many useful applications:

PROPOSITION 9.10. *Given a basis $e_1, \ldots, e_n \in V$ with dual basis $e_*^1, \ldots, e_*^n \in V^*$, we have*

$$\lambda^1 \wedge \ldots \wedge \lambda^n = \det \begin{pmatrix} \lambda^1(e_1) & \cdots & \lambda^1(e_n) \\ \vdots & \ddots & \vdots \\ \lambda^n(e_1) & \cdots & \lambda^n(e_n) \end{pmatrix} e_*^1 \wedge \ldots \wedge e_*^n$$

for any $\lambda^1, \ldots, \lambda^n \in \Lambda^1 V^*$.

PROOF. Use (9.3) to evaluate $(\lambda^1 \wedge \ldots \wedge \lambda^n)(e_1, \ldots, e_n)$, then plug in the formula (9.2) for the determinant. □

EXERCISE 9.11. Find a second proof of Proposition 9.10 using the following idea. Associate to each $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ the 1-form $\mathbf{v}_\flat := v_i e_*^i \in \Lambda^1 V^*$. What can you say about the multilinear function $\omega : \mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}$ defined by $\omega(\mathbf{v}^1, \ldots, \mathbf{v}^n) := (\mathbf{v}_\flat^1 \wedge \ldots \wedge \mathbf{v}_\flat^n)(e_1, \ldots, e_n)$?

REMARK 9.12. The formula (9.4) for the product of $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^\ell V^*$ can be written in more verbose form as

$$(9.6) \qquad (\alpha \wedge \beta)(v_1, \ldots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (-1)^{|\sigma|} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \cdot \beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}).$$

The factor in front makes this formula a bit hard to memorize, but there is a combinatorial trick that makes it easier. Let

$$S_{k,\ell} \subset S_{k+\ell}$$

denote the subset consisting of permutations σ that satisfy

$$\sigma(1) < \ldots < \sigma(k) \qquad \text{and} \qquad \sigma(k+1) < \ldots < \sigma(k+\ell);$$

such permutations are sometimes called **shuffles**. They do not form a subgroup, but every permutation in $S_{k+\ell}$ is obtained from a unique shuffle by composing it with something in the subgroup $S_k \times S_\ell \subset S_{k+\ell}$ consisting of permutations that preserve the subsets $\{1, \ldots, k\}$ and $\{k+1, \ldots, k+\ell\}$. The key observation is that there are exactly $k!\ell!$ elements in this subgroup, and applying them has the effect of permuting the sets of vectors that are plugged into each of α and β in (9.6), while simultaneously changing the sign $(-1)^{|\sigma|}$ in a way that *cancels* the resulting change in the product of α and β. The result is that (9.6) contains $k!\ell!$ times as many terms as it actually needs: it is equivalent to the simpler formula

$$(9.7) \qquad (\alpha \wedge \beta)(v_1, \ldots, v_{k+\ell}) = \sum_{\sigma \in S_{k,\ell}} (-1)^{|\sigma|} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \cdot \beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}),$$

in which no combinatorial factor is needed because the sum ranges only over shuffles.

9.3. The differential graded algebra of forms. Everything stated in the previous section implies a statement about differential forms on a manifold M, simply by replacing the vector space V with a tangent space T_pM and then letting $p \in M$ vary. In particular, a k-form $\omega \in \Omega^k(M)$ can now be understood as a function that associates to each $p \in M$ an element

$$\omega_p \in \Lambda^k T_p^* M := \Lambda^k(T_pM)^*.$$

It follows that if $\dim M = n$, then k-forms for $k > n$ are identically 0, hence the direct sum

$$\Omega^*(M) := \bigoplus_{k=0}^{\infty} \Omega^k(M)$$

has only finitely many nontrivial summands. (It is an infinite-dimensional space nonetheless, since each $\Omega^k(M)$ for $k = 0, \ldots, n$ is infinite dimensional.) The wedge product of differential forms is now defined pointwise, i.e. given $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^\ell(M)$, we define $\alpha \wedge \beta \in \Omega^{k+\ell}(M)$ by

$$(\alpha \wedge \beta)_p = \alpha_p \wedge \beta_p \in \Lambda^{k+\ell} T_p^* M.$$

The smoothness of $\alpha \wedge \beta$ by this definition will become clear momentarily when we write it down in local coordinates. Given a chart $(\mathcal{U}, x)$, the natural basis of T_pM to use at points $p \in \mathcal{U}$ is given by the coordinate vector fields $\partial_1, \ldots, \partial_n$, and its dual basis consists of the coordinate differentials $dx^1, \ldots, dx^n$. Any smooth k-form $\omega \in \Omega^k(M)$ can thus be written over $\mathcal{U}$ as

$$\begin{aligned} \omega &= \omega_{i_1 \ldots i_k}\, dx^{i_1} \otimes \ldots \otimes dx^{i_k} = \frac{1}{k!} \omega_{i_1 \ldots i_k}\, dx^{i_1} \wedge \ldots \wedge dx^{i_k} \\ &= \sum_{i_1 < \ldots < i_k} \omega_{i_1 \ldots i_k}\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}, \end{aligned} \tag{9.8}$$

where the Einstein summation convention is in effect for the first line but (in order to eliminate redundancy caused by antisymmetry) not for the second, and the smooth component functions are given by

$$\omega_{i_1 \ldots i_k} = \omega(\partial_{i_1}, \ldots, \partial_{i_k}) \in C^\infty(\mathcal{U}).$$

A coordinate formula for the wedge product can then be extracted from Exercise 9.9, namely

$$(\alpha \wedge \beta)_{i_1 \ldots i_{k+\ell}} = \frac{(k+\ell)!}{k!\ell!} \alpha_{[i_1 \ldots i_k} \beta_{i_{k+1} \ldots i_{k+\ell}]},$$

so assuming that α and β have smooth components, the same is clearly true for $\alpha \wedge \beta$. Theorem 9.4 now carries over to the statement that $\wedge$ defines a bilinear map

$$\Omega^*(M) \times \Omega^*(M) \to \Omega^*(M) : (\alpha, \beta) \mapsto \alpha \wedge \beta$$

that is associative and graded commutative, where the latter again means that for homogeneous elements $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^\ell(M)$, $\alpha \wedge \beta = \pm \beta \wedge \alpha$, with the minus sign appearing if and only if k and ℓ are both odd.

Example 9.13. Using Cartesian coordinates (x, y, z) on $\mathbb{R}^3$, the second line of (9.8) says that every $\omega \in \Omega^2(\mathbb{R}^3)$ has a unique presentation in the form

$$\omega = \omega_{xy}\, dx \wedge dy + \omega_{xz}\, dx \wedge dz + \omega_{yz}\, dy \wedge dz,$$

determined by three smooth functions $\omega_{xy}, \omega_{xz}, \omega_{yz} : \mathbb{R}^3 \to \mathbb{R}$.

Example 9.14. For $k = n$, the summation in the second line of (9.8) contains only one term. It follows that on an n-manifold M with smooth chart $(\mathcal{U}, x)$, every $\omega \in \Omega^n(M)$ can be written in local coordinates as

$$\omega = f\, dx^1 \wedge \ldots \wedge dx^n \qquad \text{on } \mathcal{U},$$

where the real-valued function $f \in C^\infty(\mathcal{U})$ is given by $f = \omega(\partial_1, \ldots, \partial_n)$.

EXERCISE 9.15. Beginners sometimes fixate on the antisymmetry of the wedge product for 1-forms and thus expect $\omega \wedge \omega = 0$ to hold always, but graded commutativity only implies this when ω has odd degree. Find a concrete example of a 2-form ω on $\mathbb{R}^4$ such that $\omega \wedge \omega \neq 0$.

We can now give a more practically useful characterization of the exterior derivative $d : \Omega^k(M) \to \Omega^{k+1}(M)$, which was defined in §8.2 via C^∞-linearity. A quick word about signs: you've already noticed that in the wedge product, a minus sign gets introduced whenever the order of two elements with odd degree is changed. One can use this same rule to remember the sign in the Leibniz rule below if one thinks of the operator d itself as an object with odd degree; it makes sense in fact to define its degree as 1, since that is the amount by which it raises the degree of any homogeneous element of $\Omega^*(M)$ fed into it.

PROPOSITION 9.16. *The exterior derivative $d : \Omega^*(M) \to \Omega^*(M)$ is the unique linear map that satisfies the following conditions:*

1. *d is* local, *meaning that for every form $\omega \in \Omega^*(M)$ and every $p \in M$, $(d\omega)_p \in \Lambda^* T_p^* M$ depends only on the restriction of ω to a neighborhood of p.*
2. *For each $f \in \Omega^0(M) = C^\infty(M)$, $df \in \Omega^1(M)$ is the differential of f.*
3. *For any homogeneous elements $\alpha, \beta \in \Omega^*(M)$, d satisfies the "graded Leibniz rule"*
$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{|\alpha|} \alpha \wedge d\beta.$$
4. *$d \circ d = 0$.*

COROLLARY 9.17. *For any chart $(\mathcal{U}, x)$ and any smooth function $f : \mathcal{U} \to \mathbb{R}$,*

$$d\left(f\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}\right) = df \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} = \partial_j f\, dx^j \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \quad \text{on } \mathcal{U}. \tag{9.9}$$

□

PROOF OF PROPOSITION 9.16. Let us start by ignoring the definition of $d : \Omega^*(M) \to \Omega^*(M)$ given in §8.2 and showing that a map satisfying the four properties stated above exists and is unique. The uniqueness follows from the observation that for any chart $(\mathcal{U}, x)$, every k-form on $\mathcal{U}$ is a sum of terms of the form $f\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}$, and if d satisfies properties (2)–(4) then its action on this particular product is given by (9.9). To prove existence, suppose first that $M = \mathcal{U}$ is the domain of a global chart x, in which case the only possible definition of d satisfying the required properties is again via (9.9). It is immediate that d by this definition satisfies properties (1) and (2); let us verify that it also satisfies (3) and (4). To prove the graded Leibniz rule, we observe first that it is true for a pair of 0-forms $f, g \in \Omega^0(M) = C^\infty(M)$, as the product rule from first-year analysis implies

$$d(fg) = df \cdot g + f \cdot dg.$$

For the general case, bilinearity allows us to restrict attention to a pair $\alpha, \beta \in \Omega^*(\mathcal{U})$ of the form $\alpha = f\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}$ and $\beta = g\, dx^{j_1} \wedge \ldots \wedge dx^{j_\ell}$. To make the notation more manageable, let us abbreviate $dx^I := dx^{i_1} \wedge \ldots \wedge dx^{i_k}$ and $dx^J := dx^{j_1} \wedge \ldots \wedge dx^{j_\ell}$; then

$$\begin{aligned} d(\alpha \wedge \beta) &= d\left(fg\, dx^I \wedge dx^J\right) = d(fg) \wedge dx^I \wedge dx^J = (df \cdot g + f \cdot dg) \wedge dx^I \wedge dx^J \\ &= \left(df \wedge dx^I\right) \wedge \left(g\, dx^J\right) + (-1)^k \left(f\, dx^I\right) \wedge \left(dg \wedge dx^J\right) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta, \end{aligned}$$

where the sign $(-1)^k$ arose when we changed the order of $dg \in \Omega^1(\mathcal{U})$ and $dx^I \in \Omega^k(\mathcal{U})$. To prove $d \circ d = 0$, we can similarly consider $\alpha = f\, dx^I$ and compute

$$d(d\alpha) = d(df \wedge dx^I) = d(\partial_j f\, dx^j \wedge dx^I) = d(\partial_j f) \wedge dx^j \wedge dx^I = \partial_k \partial_j f\, dx^k \wedge dx^j \wedge dx^I.$$

This last expression contains implied summations over both k and j, and we observe that while exchanging the roles of k and j leaves $\partial_k \partial_j f$ unchanged, it switches the sign of $dx^k \wedge dx^j$, so that every term in this sum is balanced by a cancelling term, and the sum if therefore 0.

Observe next that while our definition of $d : \Omega^*(\mathcal{U}) \to \Omega^*(\mathcal{U})$ above was expressed in terms of the specific coordinates $x^1, \dots, x^n$, the fact that it satisfies properties (1)–(4) implies that any other choice of coordinates would have given the same result, as it would also have given a definition satisfying properties (1)–(4). On a general manifold M, one can now define $d : \Omega^*(M) \to \Omega^*(M)$ on small neighborhoods using local coordinates and appeal to the fact that the definition is independent of coordinates, producing a global definition.

It remains only to prove that our definition of d via properties (1)–(4) matches the definition in §8.2. We will prove this by showing that (9.9) implies the same local coordinate formula that was derived in Proposition 8.7. Recall that in a local chart $(\mathcal{U}, x)$, an arbitrary k-form with components $\omega_{i_1 \dots i_k} = \omega(\partial_{i_1}, \dots, \partial_{i_k})$ can be written as

$$\omega = \sum_{i_1 < \dots < i_k} \omega_{i_1 \dots i_k} \, dx^{i_1} \wedge \dots \wedge dx^{i_k} = \frac{1}{k!} \omega_{i_1 \dots i_k} \, dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

where the summation convention is in effect only in the second expression, in which the combinatorial factor accounts for the fact that each term in the implied summation appears in $k!$ identical copies arising from permutations of the indices $i_1, \dots, i_k$. The formula (9.9) then implies

$$d\omega = \frac{1}{k!} d\omega_{i_1 \dots i_k} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} = \frac{1}{k!} \partial_{i_0} \omega_{i_1 \dots i_k} \, dx^{i_0} \wedge \dots \wedge dx^{i_k}.$$

In this last sum, nonzero contributions come only from terms in which the numbers $i_0, \dots, i_k \in \{1, \dots, n\}$ are all distinct, and if we write S_{k+1} for the group of bijections on $\{0, \dots, k\}$, each of these terms can be permuted by some $\sigma \in S_{k+1}$ to produce a product $dx^{i_0} \wedge \dots \wedge dx^{i_k}$ with $i_0 < \dots < i_k$, at the cost of applying the inverse permutation to the indices of $\partial_{i_0} \omega_{i_1 \dots i_k}$ and multiplying by the sign $(-1)^{|\sigma|}$. The expression therefore becomes

$$\begin{aligned} &\frac{1}{k!} \sum_{i_0 < \dots < i_k} \sum_{\sigma \in S_{k+1}} (-1)^{|\sigma|} \partial_{i_{\sigma(0)}} \omega_{i_{\sigma(1)} \dots i_{\sigma(k)}} \, dx^{i_0} \wedge \dots \wedge dx^{i_k} \\ &\quad = \frac{(k+1)!}{k!} \sum_{i_0 < \dots < i_k} \partial_{[i_0} \omega_{i_1 \dots i_k]} \, dx^{i_0} \wedge \dots \wedge dx^{i_k} = (k+1) \sum_{i_0 < \dots < i_k} \partial_{[i_0} \omega_{i_1 \dots i_k]} \, dx^{i_0} \wedge \dots \wedge dx^{i_k}, \end{aligned}$$

which matches Proposition 8.7. □

The wedge product and exterior derivative make $\Omega^*(M)$ into an example of a (commutative) **differential graded algebra** (*graduierte Differentialalgebra*), or "DGA" for short. The inclusion of the word "graded" refers in the first place to the direct sum decomposition $\Omega^*(M) = \bigoplus_{k \geqslant 0} \Omega^k(M)$, but more importantly it refers to the sign appearing in the Leibniz rule of Proposition 9.16. A similar sign prevents $\Omega^*(M)$ from satisfying the commutativity relation $\alpha \wedge \beta = \beta \wedge \alpha$ in general, but the convention is nonetheless to call it a "commutative DGA" if it satisfies the graded commutativity relation $\alpha \wedge \beta = (-1)^{|\alpha||\beta|} \beta \wedge \alpha$.

Recall from §8.3 that pullbacks of differential forms can be defined for arbitrary smooth maps $\varphi : M \to N$, not just diffeomorphisms.

PROPOSITION 9.18. *For any smooth map* $\varphi : M \to N$*:*

(1) $\varphi^*(\alpha \wedge \beta) = \varphi^*\alpha \wedge \varphi^*\beta$ *for all* $\alpha, \beta \in \Omega^*(N)$*;*
(2) $\varphi^*(d\omega) = d(\varphi^*\omega)$ *for all* $\omega \in \Omega^*(N)$*.*

PROOF. The first statement follows directly from the definitions. For the second, we start with the case $\omega = f \in C^\infty(N) = \Omega^0(N)$ and use the chain rule: $\varphi^*(df) := df \circ T\varphi = d(f \circ \varphi) =: d(\varphi^* f)$. Since every differential form is locally a finite sum of wedge products of functions and differentials, the graded Leibniz rule then extends this result to all $\omega \in \Omega^k(N)$. □

10. Oriented manifolds and the integral

10.1. Change of variables. One of the messages of the previous lecture was that on an n-manifold M, one can use differential n-forms to define sensible notions of "n-dimensional volume" and thus measures, from which a notion of integration should emerge. Let's consider first how this might work when M is an open subset $\mathcal{U} \subset \mathbb{R}^n$ in Euclidean space.

There is a canonical choice of coordinates $x^1, \dots, x^n$ on $\mathcal{U} \subset \mathbb{R}^n$, leading us naturally to consider the n-form $dx^1 \wedge \dots \wedge dx^n \in \Omega^n(\mathcal{U})$. It has the desirable property that at every point $p \in \mathcal{U}$, if one feeds into it the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ of $\mathbb{R}^n = T_p\mathcal{U}$, the result (by (9.3)) is 1, which happens also to be the Lebesgue measure of the paralelepiped spanned by these vectors, i.e. the n-dimensional unit cube. It follows that if one interprets $dx^1 \wedge \dots \wedge dx^n$ as a way of computing volumes on tangent spaces $T_p\mathcal{U} = \mathbb{R}^n$, the volume it computes is the *standard* notion of volume, i.e. the Lebesgue measure.

This observation motivates the following definition, which (in light of Example 9.14) tells us how to integrate an arbitrary compactly supported n-form on $\mathcal{U} \subset \mathbb{R}^n$.

Definition 10.1. For any integer $n \geqslant 1$, any compactly supported smooth function $f : \mathcal{U} \to \mathbb{R}$ on an open subset $\mathcal{U} \subset \mathbb{R}^n$ and any Lebesgue-measurable subset $A \subset \mathcal{U}$, the **integral** of the n-form $\omega := f\, dx^1 \wedge \dots \wedge dx^n$ over A is defined to be the Lebesgue integral of f on A with respect to the standard Lebesgue measure m on $\mathbb{R}^n$, i.e.

$$\int_A \omega = \int_A f\, dx^1 \wedge \dots \wedge dx^n := \int_A f\, dm \in \mathbb{R}.$$

Remark 10.2. If you prefer to think in terms of Riemann integrals rather than Lebesgue integrals, you are free to do so in Definition 10.1 at the cost of being slightly more restrictive about the subset $A \subset \mathcal{U}$, e.g. for almost all[37] applications it suffices to imagine that A is an open or closed subset. Nothing in our discussion of integration will depend in any serious way on the distinction between the Riemann and Lebesgue integrals. We will continue to use the language of Lebesgue integration because it seems the most natural.

Analysis conventions sometimes denote the Lebesgue measure on $\mathbb{R}^n$ more suggestively as "$dx^1 \dots dx^n$", so that Definition 10.1 becomes the easy-to-remember formula

$$\int_A f\, dx^1 \wedge \dots \wedge dx^n := \int_A f(x^1, \dots, x^n)\, dx^1 \dots dx^n.$$

Let's get a bit more ambitious now: suppose M is a more general n-manifold and $\omega \in \Omega^n(M)$ is a compactly supported top-dimensional differential form that happens to have its support contained in the domain $\mathcal{U} \subset M$ of some chart $(\mathcal{U}, x)$. In the corresponding local coordinates, ω can therefore also be written within $\mathcal{U}$ as $f\, dx^1 \wedge \dots \wedge dx^n$ for a smooth compactly supported function $f : \mathcal{U} \to \mathbb{R}$. Expressing f as a function of the coordinates $x^1, \dots, x^n$ on $\mathcal{U}$, it now seems natural to define

$$\int_A \omega := \int_{x(A)} f(x^1, \dots, x^n)\, dx^1 \dots dx^n \tag{10.1}$$

for any subset $A \subset \mathcal{U}$ such that $x(A) \subset x(\mathcal{U}) \subset \mathbb{R}^n$ is measurable, i.e. the function whose Lebesgue integral we are actually computing is $f \circ x^{-1} : x(\mathcal{U}) \to \mathbb{R}$. To see why this might be a sensible definition, write the standard Cartesian coordinates on $\mathbb{R}^n$ as $t^1, \dots, t^n$ so as to distinguish them from the coordinates $x^1, \dots, x^n$ on $\mathcal{U}$; regarding both sets of coordinates as functions on their respective domains, they are related by

$$t^i \circ x = x^i \quad \text{on } \mathcal{U}, \qquad i = 1, \dots, n. \tag{10.2}$$

[37] no pun intended

Definition 10.1 now identifies the Lebesgue integral we just described with the integral of the n-form $(f \circ x^{-1})\, dt^1 \wedge \ldots \wedge dt^n$ over $x(A) \subset x(\mathcal{U}) \subset \mathbb{R}^n$. According to Proposition 9.18 and (10.2), the diffeomorphism $M \supset \mathcal{U} \xrightarrow{x} x(\mathcal{U}) \subset \mathbb{R}^n$ pulls this n-form back to $\mathcal{U}$ as

$$x^* \left((f \circ x^{-1})\, dt^1 \wedge \ldots \wedge dt^n \right) = f \cdot \left(x^* dt^1 \wedge \ldots \wedge x^* dt^n \right) = f \cdot \left(d(x^* t^1) \wedge \ldots \wedge d(x^* t^n) \right)$$
$$= f\, dx^1 \wedge \ldots \wedge dx^n = \omega,$$

so (10.1) follows from Definition 10.1 if we stipulate that the integral should satisfy

$$\int_A x^* \alpha = \int_{x(A)} \alpha$$

for all compactly supported n-forms α on $x(\mathcal{U}) \subset \mathbb{R}^n$. This identity is consistent with our intuition about pullbacks via diffeomorphisms: x^* gives a bijection allowing geometric data on $x(\mathcal{U}) \subset \mathbb{R}^n$ to be identified with geometric data on $\mathcal{U} \subset M$, and it would make sense for our definition of the integral to respect such identifications.

But there is still a crucial question to be answered: does our definition of $\int_A \omega$ as described above depend on the choice of chart $x : \mathcal{U} \to \mathbb{R}^n$?

Suppose $y : \mathcal{U} \to \mathbb{R}^n$ is a second chart defined on the same domain, so ω can also be written as $\omega = g\, dy^1 \wedge \ldots \wedge dy^n$ for some function $g : \mathcal{U} \to \mathbb{R}$, and $\int_A \omega$ according to this chart should be $\int_{y(A)} g \circ y^{-1}\, dm$, so we need to know whether this is the same as $\int_{x(A)} f \circ x^{-1}\, dm$. To clarify this, let us abbreviate $\psi := y \circ x^{-1} : x(\mathcal{U}) \to y(\mathcal{U})$ for the transition map relating x and y, and use Proposition 9.10 to write

$$dy^1 \wedge \ldots \wedge dy^n = \det \left(\frac{\partial y}{\partial x} \right) dx^1 \wedge \ldots \wedge dx^n \qquad \text{on } \mathcal{U},$$

where we abbreviate the matrix-valued function

$$\frac{\partial y}{\partial x} := \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \cdots & \frac{\partial y^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y^n}{\partial x^1} & \cdots & \frac{\partial y^n}{\partial x^n} \end{pmatrix} : \mathcal{U} \to \mathbb{R}^{n \times n}.$$

The identity $f\, dx^1 \wedge \ldots \wedge dx^n = \omega = g\, dy^1 \wedge \ldots \wedge dy^n$ thus implies $f = g \cdot \det \left(\frac{\partial y}{\partial x} \right)$. At any point $p \in \mathcal{U}$, $\frac{\partial y}{\partial x}(p)$ is just the Jacobian matrix of the transition map ψ at $x(p)$, and this last identity thus implies

$$f \circ x^{-1} = \left(g \circ x^{-1} \right) \cdot \det D\psi.$$

If we now write $G := g \circ y^{-1}$, then $f \circ x^{-1}$ becomes $(G \circ \psi) \cdot \det D\psi$, and the identity we were hoping for becomes

$$\int_{y(A)} g \circ y^{-1}\, dm = \boxed{\int_{\psi(x(A))} G\, dm \stackrel{?}{=} \int_{x(A)} (G \circ \psi) \cdot \det D\psi\, dm} = \int_{x(A)} f \circ x^{-1}\, dm. \tag{10.3}$$

This should look familiar, as it is *almost* the classical change-of-variables formula, except for one detail: in the classical formula, the Jacobian determinant $\det(D\psi)$ is replaced by its absolute value. That is fine if $\det(D\psi)$ happens to be positive—we do of course know that it can never be 0, since ψ is a diffeomorphism and $D\psi(q) : \mathbb{R}^n \to \mathbb{R}^n$ is therefore an isomorphism for all $q \in x(\mathcal{U})$. But nothing in our discussion so far has ruled out the possibility that $\det(D\psi)$ may sometimes be *negative*, and there certainly do exist diffeomorphisms between regions in $\mathbb{R}^n$ that have negative Jacobian determinant, e.g. the reflection $(x, y) \mapsto (x, -y)$ on $\mathbb{R}^2$. The answer to the crucial question about (10.1) is therefore a resounding *sometimes*:

PROPOSITION 10.3. *In the setting of* (10.1)*, two charts defined on* $\mathcal{U}$ *give matching definitions of* $\int_A \omega$ *if the Jacobian determinant of their transition map is everywhere positive.* □

10.2. Orientations. The upshot of our change-of-variables discussion is that integrating an n-form $\omega \in \Omega^n(M)$ by writing it in local coordinates as $\omega = f\,dx^1 \wedge \ldots \wedge dx^n$ and then integrating the function f in coordinates does not give a fully coordinate-invariant result, but it will *become* coordinate-invariant if for some reason we never have to worry about transition maps whose Jacobian determinant is negative. This is our first encounter in this course with the notion of *orientation*.

DEFINITION 10.4. Given open subsets $\mathcal{U}, \mathcal{V} \subset \mathbb{R}^n$ for $n \geqslant 1$, a diffeomorphism $\psi : \mathcal{U} \to \mathcal{V}$ is called **orientation preserving** (*orientierungserhaltend*) if the Jacobian matrix $D\psi(p) \in \mathrm{GL}(n, \mathbb{R})$ at every point $p \in \mathcal{U}$ has positive determinant. It is called **orientation reversing** (*orientierungsumkehrend*) if $\det D\psi(p) < 0$ for all p.

We will say more about the intuitive meaning of this definition in a moment, but for now, you may want to keep the following linear examples in mind:

(1) Every rotation $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ defines an orientation-preserving diffeomorphism $\mathbb{R}^2 \to \mathbb{R}^2$. More generally, every element of the special orthogonal group $\mathrm{SO}(n)$ (cf. Exercise 4.25) defines an orientation-preserving diffeomorphism $\mathbb{R}^n \to \mathbb{R}^n$.
(2) The reflection $(x, y) \mapsto (x, -y)$ is an orientation-reversing diffeomorphism $\mathbb{R}^2 \to \mathbb{R}^2$, and more generally, every element of $\mathrm{O}(n)\backslash\mathrm{SO}(n)$ defines an orientation-reserving diffeomorphism $\mathbb{R}^n \to \mathbb{R}^n$. In particular, this includes every linear transformation on $\mathbb{R}^n$ that is defined by reflecting across an $(n-1)$-dimensional subspace.

DEFINITION 10.5. A smooth atlas $\mathcal{A} = \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ on a manifold M of dimension $n \geqslant 1$ is called **oriented** (*orientiert*) if all of its transition maps $x_\alpha \circ x_\beta^{-1}$ are orientation preserving. An **orientation** (*Orientierung*) of a manifold M with maximal smooth atlas $\mathcal{A}$ is a subset $\mathcal{A}^+ \subset \mathcal{A}$ that forms a maximal oriented atlas for M. A smooth manifold that has been equipped with an orientation $\mathcal{A}^+$ is called an **oriented manifold** (*orientierte Mannigfaltigkeit*), and the smooth charts in $\mathcal{A}^+$ are then called the **oriented charts**. A manifold is called **orientable** (*orientierbar*) if it admits an orientation.

One can argue as in Lemma 2.5 that given a smooth structure $\mathcal{A}$, every oriented atlas $\mathcal{A}^+ \subset \mathcal{A}$ has a unique extension to a maximal one and thus determines an orientation. In practice, we will see that there are usually more convenient ways to specify an orientation than by explicitly finding an oriented atlas, but here are a few examples where the latter can easily be done:

EXERCISE 10.6. Show that the atlas we defined on S^1 in Lecture 1 is oriented.

EXERCISE 10.7. Use the atlas from Exercise 1.7 to show that S^2 is orientable. (Depending on how you constructed the charts in that exercise, you might now have to modify them slightly for the sake of orientations.)

EXAMPLE 10.8. The manifold $\mathbb{R}^n$ carries a canonical global chart defined by the identity map, so this chart forms an oriented atlas and thus endows $\mathbb{R}^n$ with a canonical orientation.

EXAMPLE 10.9. If M has an oriented atlas $\mathcal{A}^+$ and $\mathcal{O} \subset M$ is an open subset, then the atlas $\mathcal{A}^+_{\mathcal{O}}$ on $\mathcal{O}$ constructed as in §2.4.2 is automatically also oriented, thus open subsets of oriented manifolds inherit natural orientations. In light of the previous example, this applies in particular to open subsets of $\mathbb{R}^n$.

EXERCISE 10.10. Show that if M and N are both orientable, then so is $M \times N$.

Exercise 10.11. Convince yourself that the atlases on the projective plane and Klein bottle described in §2.4.7 are not oriented. (This does not yet prove that these manifolds are not orientable, since one might imagine that there are other ways to construct an oriented atlas. But we will see below that this is impossible.)

Definition 10.12. For two oriented smooth manifolds M and N, a diffeomorphism $f : M \to N$ is called **orientation preserving** or **orientation reversing** if the map $y \circ f \circ x^{-1}$ is orientation preserving / reversing respectively for every choice of oriented smooth charts $(\mathcal{U}, x)$ on M and $(\mathcal{V}, y)$ on N.

Exercise 10.13. Show that for the orientations of S^1 and S^2 defined in Exercises 10.6 and 10.7, the antipodal map $S^n \to S^n : p \mapsto -p$ is orientation preserving for $n = 1$ but orientation reversing for $n = 2$.

Remark 10.14. In light of Definition 10.12 and the canonical orientations of $\mathbb{R}^n$ and open subsets specified by Examples 10.8 and 10.9, a smooth chart $(\mathcal{U}, x)$ on an oriented manifold M is an oriented chart if and only if the diffeomorphism $M \supset \mathcal{U} \xrightarrow{x} x(\mathcal{U}) \subset \mathbb{R}^n$ is orientation preserving.

Let's discuss next some useful alternative perspectives on the notion of orientation. We recall first the basic notion from topology of *connected components*. In topology one distinguishes between two slightly different notions of connectedness, but we will not need to worry about this distinction since for manifolds, they are equivalent.

Definition 10.15. A manifold M is **connected** (*zusammenhängend*) if for every pair of points $p, q \in M$, there exists a continuous path $\gamma : [0, 1] \to M$ with $\gamma(0) = p$ and $\gamma(1) = q$. The **connected components** (*Zusammenhangskomponenten*) of M are the maximal connected subsets.

It should be easy to convince yourself that each connected component of a manifold is both closed and open as a subset, hence it is also a manifold. In fact, if M has connected components $\{M_\alpha\}_{\alpha \in I}$, then there is a natural diffeomorphism $\coprod_{\alpha \in I} M_\alpha \cong M$.

Returning to the subject of orientations, consider a 2-dimensional subspace $P \subset \mathbb{R}^3$, i.e. a plane. One common way of characterizing what it should mean intuitively for P to be "oriented" in one way or the other is to decide which side of P is the "top" and which is the "bottom"; in other words, we draw a distinction between the two connected components of $\mathbb{R}^3 \backslash P$, labelling one component as "above" the plane and the other as "below" it. An equivalent way to say this is that one makes a choice of a unit vector $\mathbf{n} \in \mathbb{R}^3$ orthogonal to P, so that one can then decide to call the direction indicated by $\mathbf{n}$ "above" and the opposite direction "below". There are obviously two possible choices of the vector $\mathbf{n}$, and for an arbitrary plane $P \subset \mathbb{R}^3$, neither choice can be considered canonical.

Now, the case of a plane $P \subset \mathbb{R}^3$ is rather special since it is a submanifold of $\mathbb{R}^3$, and we do not want to have to assume all manifolds we consider are presented to us as submanifolds of Euclidean space. But actually, there is another way to characterize the choice of normal vector $\mathbf{n}$ in terms of vectors that are tangent to P. You may have learned it as the "right hand rule" when you first encountered vectors and the cross product in school: imagine positioning your right hand along the plane $P \subset \mathbb{R}^3$ so that your thumb points orthogonal to it in the direction of $\mathbf{n}$, but your other four fingers are tangent to P. Those four fingers will want to curl in a particular manner, defining a direction of rotation on the plane that one might choose to label "counterclockwise". (This is exactly what one does—at least in the northern hemisphere—when one visualizes the Earth "from above" and says that it rotates counterclockwise. In that situation, "from above" means that one chooses to view the Earth from a vantage point that is centered on the north pole; if one centered the picture on the south pole instead, the rotation would look clockwise! For the same reason, it is important to consistently use the *right* hand rather than the left hand when implementing the right hand rule, as switching hands would indicate a rotation in the other direction.)

The upshot of this heuristic discussion is this: our intuitive notion of what it means to orient a plane $P \subset \mathbb{R}^3$ is equivalent to making a choice of which direction of rotation on P should be labelled as counterclockwise instead of clockwise. This notion can be defined on *any* surface Σ by talking about rotations in the tangent spaces $T_p\Sigma$, and there is no longer any need to discuss normal vectors or assume an embedding $\Sigma \hookrightarrow \mathbb{R}^3$ is given. Moreover, we will see presently that instead of specifying a preferred direction of rotation in $T_p\Sigma$, it is equivalent to specify a preferred class of ordered bases.

DEFINITION 10.16. For a vector space V of dimension $n \geqslant 1$, let

$$\mathcal{B}(V) \subset V^{\times n} := \underbrace{V \times \ldots \times V}_{n}$$

denote the set of all ordered n-tuples $(v_1, \ldots, v_n)$ that form bases of V.

Observe that $\mathcal{B}(V)$ is an open subset of $V^{\times n}$ since linear independence cannot be destroyed by small perturbations. In fact, after choosing any isomorphism $V \to \mathbb{R}^n$, the vectors in any tuple $(v_1, \ldots, v_n) \in \mathcal{B}(V)$ can be put together as columns of an n-by-n matrix, thus identifying $\mathcal{B}(V)$ with the general linear group $\mathrm{GL}(n, \mathbb{R})$, which is indeed an open subset of the space of matrices $\mathbb{R}^{n\times n}$.

Now consider the case $V = \mathbb{R}^2$. Given any $(v_1, v_2) \in \mathcal{B}(\mathbb{R}^2)$, moving from the direction of v_1 to that of v_2 requires a rotation of less than 180 degrees that is either counterclockwise or clockwise; for example, a counterclockwise rotation is required in order to move from the first standard basis vector $e_1 = (1, 0)$ to the second one $e_2 = (0, 1)$, but if we exchange their roles and order the standard basis as $(e_2, e_1) \in \mathcal{B}(\mathbb{R}^2)$, then getting from e_2 to e_1 requires a clockwise rotation. For a tangent space $T_p\Sigma$ to a surface Σ, the implication is that if one has chosen which rotations to call counterclockwise as opposed to clockwise, then one has also chosen a preferred class of ordered bases $(X_1, X_2) \in \mathcal{B}(T_p\Sigma)$, i.e. we call (X_1, X_2) a ***positively oriented*** basis of the rotation moving from X_1 to X_2 is counterclockwise, and ***negatively oriented*** if that rotation is clockwise. The following facts should now be apparent:

(1) If $(X_1, X_2) \in \mathcal{B}(T_p\Sigma)$ is positively oriented, then every $(X_1', X_2') \in \mathcal{B}(T_p\Sigma)$ that can be connected to (X_1, X_2) by a continuous path in $\mathcal{B}(T_p\Sigma)$ is also positively oriented. Conversely, any two choices of positively oriented basis are related to each other by a continuous deformation of ordered bases, meaning they are connected by a continuous path in $\mathcal{B}(T_p\Sigma)$. Both statements also apply of course to negatively oriented bases.
(2) Any choice of basis $(X_1, X_2) \in \mathcal{B}(T_p\Sigma)$ can be used to *define* the distinction between clockwise and counterclockwise rotation in $T_p\Sigma$: one simply chooses it so that (X_1, X_2) is a positively oriented basis.
(3) An ordered basis (X_1, X_2) is positively oriented if and only if (X_2, X_1) is negatively oriented.

There is a basic fact about $\mathrm{GL}(2, \mathbb{R})$ in the background of the first observation above: it has exactly two connected components, characterized by the conditions $\det(\mathbf{A}) > 0$ and $\det(\mathbf{A}) < 0$. This turns out to be true in every dimension:

PROPOSITION 10.17. *For every $n \in \mathbb{N}$, the sets of* $\mathrm{GL}_+(n, \mathbb{R}) := \{\mathbf{A} \in \mathrm{GL}(n, \mathbb{R}) \mid \det(\mathbf{A}) > 0\}$ *and* $\mathrm{GL}_-(n, \mathbb{R}) := \{\mathbf{A} \in \mathrm{GL}(n, \mathbb{R}) \mid \det(\mathbf{A}) < 0\}$ *are both connected.*

PROOF. Since $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$, it suffices to prove that $\mathrm{GL}_+(n, \mathbb{R})$ is connected. To start with, we use polar decomposition to reduce this to a statement about the special orthogonal group $\mathrm{SO}(n)$. Given $\mathbf{A} \in \mathrm{GL}_+(n, \mathbb{R})$, the matrix $\mathbf{A}^T\mathbf{A}$ is symmetric and positive definite, thus it is diagonalizable with only positive eigenvalues, and therefore admits a "square root"

$$\mathbf{P} := \sqrt{\mathbf{A}^T\mathbf{A}},$$

defined in the same orthogonal basis by taking the square roots of the eigenvalues. Clearly $\mathbf{P}$ is also symmetric and positive definite, and it is now straightforward to check that $\mathbf{R} := \mathbf{A}\mathbf{P}^{-1}$ satisfies $\mathbf{R}^T\mathbf{R} = \mathbb{1}$, i.e. it is orthogonal; moreover, $\mathbf{R} \in \mathrm{SO}(n)$ since $\mathbf{A}$ and $\mathbf{P}^{-1}$ each have positive determinant. Now choose a continuous path of symmetric positive-definite matrices $\{\mathbf{P}_t\}_{t\in[0,1]}$ such that $\mathbf{P}_1 = \mathbf{P}$ and $\mathbf{P}_0 = \mathbb{1}$; such a path can be found by fixing the orthonormal eigenbasis of $\mathbf{P}$ while deforming all its (positive!) eigenvalues to 1. The path $\mathbf{A}_t := \mathbf{R}\mathbf{P}_t$ then connects $\mathbf{A}_1 = \mathbf{A}$ to $\mathbf{A}_0 = \mathbf{R} \in \mathrm{SO}(n)$, so we will be done if we can show that $\mathrm{SO}(n)$ is connected.

We argue the latter by induction: the case $n = 1$ is already clear since $\mathrm{SO}(1) = \{1\}$. Assuming $\mathrm{SO}(n-1)$ is already known to be connected, suppose $\mathbf{A} \in \mathrm{SO}(n)$ is given. We claim that there exists a continuous path $\{\mathbf{A}_t \in \mathrm{SO}(n)\}_{t\in[0,1]}$ such that $\mathbf{A}_1 = \mathbf{A}$ and $\mathbf{A}_0$ is a matrix of the form

$$\mathbf{A}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{B} \end{pmatrix}, \qquad \text{for some } \mathbf{B} \in \mathrm{SO}(n-1).$$

Observe that this claim implies the inductive step, as $\mathrm{SO}(n-1)$ is already known to be connected. To prove the claim, first choose any continuous path of unit vectors $v_1(t) \in \mathbb{R}^n$ such that $v_1(1)$ is the first column of $\mathbf{A}$ and $v_1(0)$ is the first standard basis vector $e_1 = (1, 0, \dots, 0)$; this is possible since the unit sphere S^{n-1} is connected. For any $t_0 \in [0,1]$, one can complete $v_1(t_0)$ to an orthonormal basis $v_1(t_0), \dots, v_n(t_0) \in \mathbb{R}^n$, and then find a connected neighborhood $J \subset [0,1]$ of t_0 such that the set of vectors $v_1(t), v_2(t_0), \dots, v_n(t_0)$ remains linearly independent for every $t \in J$. Now define a continuous family of orthonormal bases $v_1(t), v_2(t), \dots, v_n(t)$ for $t \in J$ by applying the Gram-Schmidt algorithm to $v_1(t), v_2(t_0), \dots, v_n(t_0)$; regarding these as columns of a matrix, we have in this way constructed a continuous family of orthogonal matrices $\{\widehat{\mathbf{A}}_t \in \mathrm{O}(n)\}_{t\in J}$ whose first columns are $v_1(t)$. Their determinants depend continuously on t and are thus either $+1$ or -1 for all $t \in J$; in the latter case, we can replace $v_n(t)$ by $-v_n(t)$ in order to assume $\widehat{\mathbf{A}}_t \in \mathrm{SO}(n)$ without loss of generality. Since $[0,1]$ is compact, we can cover it with finitely many neighborhoods J as described above, and in this way construct a family of matrices $\{\widehat{\mathbf{A}}_t \in \mathrm{SO}(n)\}_{t\in[0,1]}$ that satisfy $\mathbf{A}_1 = \mathbf{A}$ and $\mathbf{A}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{B} \end{pmatrix}$, and such that the first column of $\mathbf{A}_t$ depends continuously on t, while the other columns are continuous except at finitely many points $0 < t_1 < \dots t_N < 1$, where there are jump discontinuities. At any of these points t_j, the two matrices

$$\widehat{\mathbf{A}}_{t_j}^- := \lim_{t\to t_j^-} \widehat{\mathbf{A}}_t, \qquad \widehat{\mathbf{A}}_{t_j}^+ := \lim_{t\to t_j^+} \widehat{\mathbf{A}}_t$$

may differ, but they have the same first column, namely $v_1(t_j)$. But expressing these matrices in any orthonormal basis that starts with $v_1(t_j)$ puts both of them in the form $\begin{pmatrix} 1 & 0 \\ 0 & \mathbf{B}_\pm \end{pmatrix}$ for some $\mathbf{B}_\pm \in \mathrm{SO}(n-1)$, and by the inductive hypothesis, there exists a continuous path in $\mathrm{SO}(n-1)$ from $\mathbf{B}_-$ to $\mathbf{B}_+$. In this way, we can insert extra intervals at each of the points t_j and fill in the discontinuities, then reparametrize the interval to construct the continuous family $\mathbf{A}_t$ in the claim. □

COROLLARY 10.18. *For any vector space V of dimension $n \geqslant 1$, the set of ordered bases $\mathcal{B}(V)$ has exactly two connected components.* □

REMARK 10.19. It is very important in this entire discussion that we are talking about *real* vector spaces, not complex. In particular, the analogous set of ordered complex bases on a complex vector space is *connected*, due to the fact that $\mathrm{GL}(n,\mathbb{C})$ is connected. A hint of this is provided by the fact that the determinant on $\mathrm{GL}(n,\mathbb{C})$ takes values in $\mathbb{C}\backslash\{0\}$, which is connected, unlike $\mathbb{R}\backslash\{0\}$. As a consequence, there is no meaningful notion of orientations for *complex* manifolds; actually, every complex manifold can also be regarded as a real manifold and is orientable as a real manifold, but the orientation is *canonically* determined by its complex structure. The reason for the latter is that if we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ via the correspondence $\mathbb{C}^n \ni \mathbf{x}+i\mathbf{y} \leftrightarrow (\mathbf{x},\mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$, then every complex-linear isomorphism $\mathbf{A} \in \mathrm{GL}(n,\mathbb{C})$ becomes an element of $\mathrm{GL}(2n,\mathbb{R})$ with *positive* determinant.

Exercise 10.20 (just for fun). Adapt the proof of Proposition 10.17 to prove that $\mathrm{GL}(n,\mathbb{C})$ is connected for every $n \in \mathbb{N}$.
Hint: $\mathrm{O}(1)$ is not connected, but $\mathrm{U}(1)$ is.

We can now give a general definition of orientations of vector spaces and relate it to the previously defined notion of oriented manifolds.

Definition 10.21. An **orientation** $\mathfrak{o}_V$ of an n-dimensional vector space V for $n \geqslant 1$ is a labelling of the two connected components of $\mathcal{B}(V)$ as $\mathcal{B}^+(V)$ and $\mathcal{B}^-(V)$, which are then said to consist of the **positively oriented** and **negatively oriented** bases respectively. An **oriented vector space** is a vector space that has been equipped with an orientation. A linear isomorphism $A : V \to W$ between two oriented vector spaces is called **orientation preserving** if for every positively-oriented basis $(v_1,\ldots,v_n)$ of V, $(Av_1,\ldots,Av_n)$ is a positively-oriented basis of W, and A is otherwise called **orientation reversing**.

Notice that unlike manifolds, vector spaces always admit orientations, and there are always exactly two possible choices of orientation.

Example 10.22. As a vector space, $\mathbb{R}^n$ carries a canonical orientation for which the standard basis is regarded as positively oriented.

Exercise 10.23. Show that for the vector space $\mathbb{R}^n$ with its canonical orientation, an invertible linear map $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ is orientation preserving if and only if $\det(\mathbf{A}) > 0$.
Hint: The identity map $\mathbb{R}^n \to \mathbb{R}^n$ is clearly orientation preserving.

In light of Exercise 10.23, a diffeomorphism $\psi : \mathcal{U} \to \mathcal{V}$ between two open subsets $\mathcal{U},\mathcal{V} \subset \mathbb{R}^n$ is orientation preserving as in Definition 10.4 if and only if its derivative at every point is an orientation-preserving isomorphism $\mathbb{R}^n \to \mathbb{R}^n$ in the sense of Definition 10.21. We only need one more notion before we can set up a precise correspondence between orientations of manifolds and of their tangent spaces:

Definition 10.24. Suppose M is an n-manifold with $n \geqslant 1$, P is a topological space, $\phi : P \to M$ is a continuous map, and we consider the family of tangent spaces $\{T_{\phi(s)}M\}_{s\in P}$ at points parametrized by the map ϕ. A **continuous family of orientations along** $\phi : P \to M$ is a family $\{\mathfrak{o}_s\}_{s\in P}$, where $\mathfrak{o}_s$ is an orientation of $T_{\phi(s)}M$ for each $s \in P$, such that for every $s_0 \in P$, there exists a neighborhood $\mathcal{O} \subset P$ of s_0 and a collection of continuous maps $X_1,\ldots,X_n : \mathcal{O} \to TM$ for which $(X_1(s),\ldots,X_n(s))$ is a positively-oriented basis of $T_{\phi(s)}M$ with respect to $\mathfrak{o}_s$ for each $s \in \mathcal{O}$. In the case $P = M$ with ϕ chosen to be the identity map, we will simply refer to this as a **continuous family of orientations of the tangent spaces** of M.

Proposition 10.25. *On smooth manifolds M of dimension $n \geqslant 1$, there is a natural bijective correspondence between orientations of M and continuous families of orientations of the tangent spaces of M, and it is uniquely determined by the condition that for any diffeomorphism $f : M \to N$ between two smooth oriented manifolds, f is orientation preserving if and only if the isomorphism $T_pf : T_pM \to T_{f(p)}N$ is orientation preserving for every $p \in M$. Equivalently, a chart $(\mathcal{U},x)$ is* *oriented if and only if the corresponding basis of coordinate vector fields* $(\partial_1, \dots, \partial_n)$ *is positively oriented for every* $p \in \mathcal{U}$.

Proof. If M is oriented, one defines the orientation of T_pM for any $p \in M$ such that for any oriented chart $(\mathcal{U}, x)$ with $p \in \mathcal{U}$, the isomorphism $d_px : T_pM \to \mathbb{R}^n$ is orientation preserving (for the canonical orientation of $\mathbb{R}^n$). This is equivalent to the condition stated above involving coordinate vector fields, and the definition is independent of the choice of oriented chart since if $(\mathcal{V}, y)$ is a different choice, then d_py is the composition of d_px with an isomorphism $\mathbb{R}^n \to \mathbb{R}^n$ (defined by differentiating a transition map) that is orientation preserving. Conversely, given a continuous family of orientations of the tangent spaces T_pM, one defines the corresponding orientation of M such that a chart $(\mathcal{U}, x)$ is oriented if and only if $d_px : T_px \to \mathbb{R}^n$ is orientation preserving for every $p \in \mathcal{U}$. We leave it as an exercise to check that these definitions satisfy all of the stated properties. □

The fact that the orientations of the tangent spaces T_pM vary *continuously* with p is crucial, and it provides the easiest means of proving statements about orientations in many concrete examples.

Exercise 10.26. For a smooth n-manifold M with $n \geqslant 1$, prove:

(1) If M is connected and orientable, then it admits exactly two choices of orientation.
(2) M is orientable if and only if for every continuous path $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1)$ and every continuous family of orientations $\{\mathfrak{o}_t\}_{t\in[0,1]}$ along γ, $\mathfrak{o}_0 = \mathfrak{o}_1$.

Exercise 10.27. Show that S^n is orientable for every $n \in \mathbb{N}$.
Hint: For every $p \in S^n$ *and any basis* $X_1, \dots, X_n$ *of* T_pS^n, $(X_1, \dots, X_n, p)$ *forms a basis of* $\mathbb{R}^{n+1}$. *Use the fact that* $\mathbb{R}^{n+1}$ *is orientable.*

Exercise 10.28. Use Exercise 10.26 to show that the projective plane $\mathbb{RP}^2$ and the Klein bottle are not orientable.

Example 10.29. The physical universe is a 3-manifold, as you can plainly see by looking around you; from your local perspective it looks like $\mathbb{R}^3$, but since you cannot see the whole thing, it could in theory be diffeomorphic to any 3-manifold, even one that is not orientable. If indeed it is not orientable, then it is possible in theory for an astronaut to return from a long journey through space and find that what she used to call her right hand is now on the left side, and vice versa. She would not see it that way since her right and left eyes would also have been interchanged, but she would think that all writing now appears backwards, and the Earth (when viewed from the north pole) is now rotating clockwise. I am not aware of any law of physics that would rule out this scenario.

10.3. Definition of the integral. We are now in a position to define the integral of a compactly supported n-form on an oriented n-manifold for each $n \geqslant 1$. Denote the **support** (*Träger*) of a k-form $\omega \in \Omega^k(M)$ by

$$\operatorname{supp}(\omega) := \overline{\{p \in M \mid \omega_p \neq 0\}} \subset M,$$

and define the vector space

$$\Omega^k_c(M) := \{\omega \in \Omega^k(M) \mid \operatorname{supp}(\omega) \subset M \text{ is compact}\} \subset \Omega^k(M).$$

In the most interesting examples for our purposes, M will often be a compact manifold, in which case $\Omega^k_c(M) = \Omega^k(M)$. We will call a subset $A \subset M$ **measurable** if for every smooth chart $(\mathcal{U}, x)$ on M, the set $x(\mathcal{U} \cap A) \subset \mathbb{R}^n$ is Lebesgue measurable. The following theorem serves simultaneously as a definition.

Theorem 10.30. *For $n \in \mathbb{N}$, one can uniquely associate to every smooth oriented n-manifold M and measurable subset $A \subset M$ a linear map*

$$\Omega_c^n(M) \to \mathbb{R} : \omega \mapsto \int_A \omega$$

such that the following conditions are satisfied:

(1) If $\mathcal{U} \subset M$ is an open subset containing $\operatorname{supp}(\omega) \cap A$*, then then $\int_{\mathcal{U}\cap A} \omega = \int_A \omega$.*
(2) For $M = \mathcal{U} \subset \mathbb{R}^n$ an open subset of Euclidean space with its canonical orientation and the standard Cartesian coordinates $x^1, \dots, x^n$,

$$\int_A f\, dx^1 \wedge \dots \wedge dx^n = \int_A f\, dm$$

for all smooth compactly supported functions $f : \mathcal{U} \to \mathbb{R}$, where the right hand side is the standard Lebesgue integral of f.
(3) For any orientation-preserving diffeomorphism $\psi : M \to N$ between a pair of oriented n-manifolds,

$$\int_A \psi^*\omega = \int_{\psi(A)} \omega$$

holds for all $\omega \in \Omega_c^n(N)$ and measurable subsets $A \subset M$.

To summarize, the integral on arbitrary oriented manifolds is uniquely determined by its definition on open subsets of $\mathbb{R}^n$ and the change-of-variables formula, which now appears as the condition that integrals are invariant under pullbacks via orientation-preserving diffeomorphisms. We will prove this in the next lecture, but it is already easy to explain the idea. For forms $\omega \in \Omega_c^n(M)$ with $\operatorname{supp}(\omega)$ contained in the domain of a single oriented chart $(\mathcal{U}, x)$, one can write

$$\omega = f\, dx^1 \wedge \dots \wedge dx^n = x^*\left((f \circ x^{-1})\, dt^1 \wedge \dots \wedge dt^n\right) \qquad \text{on } \mathcal{U}$$

in terms of the standard Cartesian coordinates $t^1, \dots, t^n$ on $x(\mathcal{U}) \subset \mathbb{R}^n$ and a uniquely determined function $f : \mathcal{U} \to \mathbb{R}$. The three properties in the statement above then reproduce the definition of $\int_A \omega$ that we saw in §10.1, namely

$$\int_A \omega = \int_{\mathcal{U}\cap A} \omega = \int_{\mathcal{U}\cap A} x^*((f \circ x^{-1})\, dt^1 \wedge \dots \wedge dt^n) = \int_{x(\mathcal{U}\cap A)} (f \circ x^{-1})\, dt^1 \wedge \dots \wedge dt^n$$
$$= \int_{x(\mathcal{U}\cap A)} f \circ x^{-1}\, dm.$$

The restriction to oriented charts guarantees moreover in light of Proposition 10.3 that this result does not depend on the choice of the chart $(\mathcal{U}, x)$, though it does depend on the orientation. Linearity will then determine $\int_A \omega$ uniquely for every $\omega \in \Omega_c^n(M)$ if we can be assured that every such form is a finite sum of forms that each have compact support in the domain of some oriented chart. This is true, but not completely obvious—it will require a brief digression on the topic of *partitions of unity*, which will have many further uses as we move forward.

11. Integration and volume

11.1. Existence of the integral. I owe you a proof of Theorem 10.30 on the existence and properties of the linear map $\Omega_c^n \to \mathbb{R} : \omega \mapsto \int_A \omega$ for all oriented n-manifolds M and measurable subsets $A \subset M$. The following will serve as a useful tool for "localizing" such constructions:

Lemma 11.1. *Given a smooth manifold M, a compact subset $K \subset M$ and a finite collection of open sets $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ that cover K, there exists a collection of smooth functions $\{\varphi_\alpha : M \to [0, 1]\}_{\alpha\in I}$ satisfying the following two conditions:*

(1) For each $\alpha \in I$, φ_α has compact support contained in $\mathcal{U}_\alpha$;
(2) $\sum_{\alpha\in I} \varphi_\alpha \equiv 1$ on K.

PROOF. For each $p \in K$, choose any $\alpha_p \in I$ such that $p \in \mathcal{U}_{\alpha_p}$, and choose also a smooth function $\psi_p : M \to [0,1]$ with compact support in $\mathcal{U}_{\alpha_p}$ such that $\psi_p > 0$ on some open neighborhood $\mathcal{V}_p \subset \mathcal{U}_{\alpha_p}$ of p. The sets $\{\mathcal{V}_p\}_{p\in K}$ then form an open cover of the compact set K and therefore admit a finite subcover, i.e. there is a finite subset $K_0 \subset K$ such that $K \subset \bigcup_{p\in K_0} \mathcal{V}_p$. Now for each $\alpha \in I$, define a smooth function $\psi_\alpha : M \to [0,\infty)$ by

$$\psi_\alpha := \sum_{\{p\in K_0 \mid \alpha_p = \alpha\}} \psi_p.$$

By construction, ψ_α has compact support in $\mathcal{U}_\alpha$, and for each $q \in K$, there exists $p \in K_0$ such that $q \in \mathcal{V}_p$ and thus $\psi_p(q) > 0$, implying $\psi_{\alpha_p}(q) > 0$. It follows that $\sum_{\alpha\in I} \psi_\alpha > 0$ everywhere on K, and therefore also on some neighborhood $\mathcal{V} \subset M$ of K. On the neighborhood $\mathcal{V}$, we define

$$\varphi_\alpha := \frac{\psi_\alpha}{\sum_{\beta\in I} \psi_\beta}, \qquad \text{for each } \alpha \in I,$$

so that each φ_α takes values in $[0,1]$ and $\sum_{\alpha\in I} \varphi_\alpha \equiv 1$ by construction. Now choose any smooth function $f : M \to [0,1]$ that equals 1 on K and has compact support in $\mathcal{V}$, modify each φ_α by multiplying it by f, and extend the modified function to the rest of M so that it vanishes outside of $\mathcal{V}$. □

The collection of functions $\{\varphi_\alpha\}_{\alpha\in I}$ in this lemma is a special case of a general construction called a **partition of unity** subordinate to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ (*eine der Überdeckung untergeordnete Zerlegung der Eins*). We will extend this notion later, when we discuss more general existence theorems for geometric structures such as Riemannian metrics.

PROOF OF THEOREM 10.30. Given an oriented n-manifold M with measurable subset $A \subset M$ and $\omega \in \Omega^n_c(M)$, choose an open subset $M_0 \subset M$ that contains $\operatorname{supp}(\omega) \cap A$ but has compact closure $\overline{M}_0 \subset M$. By compactness, we can cover $\overline{M}_0$ with a finite collection of open sets $\{\mathcal{U}_\alpha \subset M\}_{\alpha\in I}$ that are domains of oriented charts $(\mathcal{U}_\alpha, x_\alpha)$, and Lemma 11.1 provides a partition of unity $\{\varphi_\alpha : M \to [0,1]\}_{\alpha\in I}$ such that

(i) φ_α has compact support contained in $\mathcal{U}_\alpha$ for each $\alpha \in I$;
(ii) $\sum_{\alpha\in I} \varphi_\alpha \equiv 1$ on M_0.

We can now write

$$\omega = \sum_{\alpha\in I} \varphi_\alpha \omega \qquad \text{on } M_0,$$

and observe that $\varphi_\alpha\omega \in \Omega^n_c(\mathcal{U}_\alpha)$, so if the integral satisfies the properties stated in the theorem, then

$$\int_A \omega = \int_{M_0\cap A} \omega = \sum_{\alpha\in I} \int_{M_0\cap A} \varphi_\alpha\omega = \sum_{\alpha\in I} \int_{\mathcal{U}_\alpha\cap A} \varphi_\alpha\omega = \sum_{\alpha\in I} \int_{x_\alpha(\mathcal{U}_\alpha\cap A)} f_\alpha \, dm, \tag{11.1}$$

where $f_\alpha : x_\alpha(\mathcal{U}_\alpha) \to \mathbb{R}$ is the unique function such that $\varphi_\alpha\omega = x^*_\alpha(f_\alpha \, dx^1 \wedge \ldots \wedge dx^n)$ on $\mathcal{U}_\alpha$. This specifies the integral uniquely.

We claim next that if $\int_A \omega \in \mathbb{R}$ is defined via (11.1), then the result is independent of all choices, namely the open subset $M_0 \subset M$ containing $\operatorname{supp}(\omega) \cap A$, the finite collection of oriented charts $\{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha\in I}$ and the functions $\{\varphi_\alpha\}_{\alpha\in I}$ satisfying (i) and (ii) above. Independence of the choice of charts follows from the discussion in §10.1, in particular Proposition 10.3. This is the step at which it is crucial that M comes with an orientation, so the transition maps that we feed into Proposition 10.3 are all orientation preserving. With this out of the way, suppose

$\{(\mathcal{V}_\beta, y_\beta)\}_{\beta \in J}$ is another finite collection of oriented charts and $\{\psi_\beta : M \to [0,1]\}_{\beta \in J}$ a collection of smooth functions that each have compact support in the corresponding subsets $\mathcal{V}_\beta$ and satisfy $\sum_{\beta \in J} \psi_\beta \equiv 1$ on some open set $M_1 \subset M$ containing $\operatorname{supp}(\omega) \cap A$. The open set $M_0 \cap M_1 \subset M$ then also contains $\operatorname{supp}(\omega) \cap A$, and is covered by the finite collection of open sets

$$\{\mathcal{U}_\alpha \cap \mathcal{V}_\beta\}_{(\alpha,\beta) \in I \times J},$$

with the functions $\{\varphi_\alpha \psi_\beta : M \to [0,1]\}_{(\alpha,\beta) \in I \times J}$ having compact support in $\mathcal{U}_\alpha \cap \mathcal{V}_\beta$ and satisfying $\sum_{(\alpha,\beta) \in I \times J} \varphi_\alpha \psi_\beta \equiv 1$ on $M_0 \cap M_1$. Any oriented chart x_α defined on $\mathcal{U}_\alpha$ is also defined on $\mathcal{U}_\alpha \cap \mathcal{V}_\beta$ for each $\beta \in J$, so we can use it to compute $\int_{\mathcal{U}_\alpha \cap \mathcal{V}_\beta \cap A} \varphi_\alpha \psi_\beta \omega$ as a Lebesgue integral over $x_\alpha(\mathcal{U}_\alpha \cap A) \subset \mathbb{R}^n$ of a function with compact support in the region $x_\alpha(\mathcal{U}_\alpha \cap \mathcal{V}_\beta)$, and the additivity of the Legesgue integral then implies

$$\int_{\mathcal{U}_\alpha \cap A} \varphi_\alpha \omega = \sum_{\beta \in J} \int_{\mathcal{U}_\alpha \cap \mathcal{V}_\beta \cap A} \varphi_\alpha \psi_\beta \omega,$$

and therefore also

$$\sum_{\alpha \in I} \int_{\mathcal{U}_\alpha \cap A} \varphi_\alpha \omega = \sum_{(\alpha,\beta) \in I \times J} \int_{\mathcal{U}_\alpha \cap \mathcal{V}_\beta \cap A} \varphi_\alpha \psi_\beta \omega.$$

But if we carry out the same argument instead with the charts $(\mathcal{V}_\beta, y_\beta)$ and write $\psi_\beta \omega = \sum_{\alpha \in I} \varphi_\alpha \psi_\beta \omega$, we find that the right hand side is also equal to $\sum_{\beta \in J} \int_{\mathcal{V}_\beta \cap A} \psi_\beta \omega$, proving that the two definitions of $\int_A \omega$ obtained from these different partitions of unity match.

It remains to check that our general definition of $\int_A \omega$ satisfies the three properties stated in the theorem, but this is easy, so we will leave it as an exercise with the following hint: the freedom to choose any convenient collection of oriented charts makes the formula $\int_A \psi^* \omega = \int_{\psi(A)} \omega$ for orientation-preserving diffeomorphisms $\psi : M \to N$ virtually a tautology. □

11.2. Computational tools. The notion of integration defined in Theorem 10.30 has several useful properties that were not mentioned yet, some of which can be applied to make actual calculations considerably easier, e.g. it is rarely actually necessary in practice to choose a partition of unity. We begin with two properties whose proofs are easy exercises.

Exercise 11.2. Prove that for an oriented n-manifold M and $\omega \in \Omega^n_c(M)$, the following properties hold:

1. If $A, B \subset M$ are two disjoint measurable subsets, then $\int_{A \cup B} \omega = \int_A \omega + \int_B \omega$.
2. If $A \subset M$ has the property that $x(\mathcal{U} \cap A) \subset \mathbb{R}^n$ has Lebesgue measure zero[38] for all smooth charts $(\mathcal{U}, x)$, then $\int_A \omega = 0$.

One frequently occurring situation in simple examples is that the domain $A \subset M$ where we want to integrate lies almost entirely inside the domain of a single chart, where the word "almost" in this case carries its usual measure-theoretic meaning, i.e. "outside of a set of measure zero". In combination with the exercise above, the next result will then allow us to dispense entirely with partitions of unity and compute the integral in a single chart:

Proposition 11.3. *Suppose M is an oriented n-manifold and $(\mathcal{U}, x)$ is an oriented chart on M. Then for any measurable subset $A \subset \mathcal{U}$ and $\omega \in \Omega^n_c(M)$ taking the form $f\, dx^1 \wedge \ldots \wedge dx^n$ in $\mathcal{U}$, the function $f \circ x^{-1}$ is Lebesgue integrable on $x(A) \subset \mathbb{R}^n$ and*

$$\int_A \omega = \int_{x(A)} f \circ x^{-1}\, dm.$$

[38] We say in this case that $A \subset M$ has **measure zero**. Note that it is not actually necessary to define a measure on M in order to define this notion.

Proof. Let $K \subset M$ denote the closure of $\mathrm{supp}(\omega) \cap A \subset M$, and observe that this set is compact since it is a closed subset of $\mathrm{supp}(\omega)$, and it is also contained in the closure of $\mathcal{U}$ since $A \subset \mathcal{U}$. In particular, the set

$$\partial K := K \cap (M \backslash \mathcal{U})$$

is contained in the boundary of the closure of $\mathcal{U}$, and by assumption it is disjoint from A. Next choose a finite collection of oriented charts $\{(\mathcal{O}_\alpha, x_\alpha)\}_{\alpha \in I}$ such that

$$K \subset \mathcal{U} \cup \bigcup_{\alpha \in I} \mathcal{O}_\alpha,$$

and for each $N \in \mathbb{N}$ and $\alpha \in I$, let

$$\mathcal{O}_\alpha^N := \left\{ p \in \mathcal{O} \;\middle|\; |x_\alpha(p) - x_\alpha(q)| < 1/N \text{ for some } q \in \partial K \cap \mathcal{O}_\alpha \right\}.$$

We observe the following:

(1) $K \subset \mathcal{U} \cup \bigcup_{\alpha \in I} \mathcal{O}_\alpha^N$ for every $N \in \mathbb{N}$.

(2) For each $\alpha \in I$, $\mathcal{O}_\alpha^1 \supset \mathcal{O}_\alpha^2 \supset \mathcal{O}_\alpha^3 \supset \ldots$, and, since $A \cap \partial K = \varnothing$,

$$A \cap \bigcap_{N \in \mathbb{N}} \mathcal{O}_\alpha^N = \varnothing. \tag{11.2}$$

For each $N \in \mathbb{N}$, we can choose a partition of unity consisting of functions $\varphi^N, \varphi_\alpha^N : M \to [0,1]$ for each $\alpha \in I$ with compact supports $\mathrm{supp}(\varphi^N) \subset \mathcal{U}$ and $\mathrm{supp}(\varphi_\alpha^N) \subset \mathcal{O}_\alpha^N$ such that $\varphi^N + \sum_{\alpha \in I} \varphi_\alpha^N \equiv 1$ on K. Since K contains $A \cap \mathrm{supp}(\omega)$, we then have

$$\int_A \omega = \int_A \varphi^N \omega + \sum_{\alpha \in I} \varphi_\alpha^N \omega$$

for every $N \in \mathbb{N}$. But for each $\alpha \in I$, (11.2) implies that the Lebesgue measure of $x_\alpha(\mathcal{O}_\alpha^N \cap A)$ converges to 0 as $N \to \infty$, thus

$$\lim_{N \to \infty} \int_A \varphi_\alpha^N \omega = 0,$$

from which follows

$$\int_A \varphi^N \omega \to \int_A \omega \qquad \text{as } N \to \infty.$$

Writing $\omega = x^*(f\, dx^1 \wedge \ldots \wedge dx^n)$ on $\mathcal{U}$ for a suitable function $f : x(\mathcal{U}) \to \mathbb{R}$, $\int_A \varphi^N \omega$ becomes the Lebesgue integral

$$\int_{x(A)} (\varphi^N \circ x^{-1}) f\, dm,$$

in which the integrand converges pointwise to f since each point in A is outside the support of all the φ_α^N for N sufficiently large. If you already believe that f is Lebesgue integrable on $x(A)$, then since $|(\varphi^N \circ x^{-1}) f| \leqslant |f|$, the dominated convergence theorem now implies that this integral converges to $\int_{x(A)} f\, dm$ as $N \to \infty$, and the latter is therefore $\int_A \omega$.

Here is a quick sketch of the proof that f really is Lebesgue integrable on $x(A)$: suppose ω is replaced by a *continuous* n-form $|\omega|$ on M that equals $-\omega$ at any point where ω evaluates negatively on some positive basis, but is otherwise identical to ω. In general $|\omega|$ will not be smooth—just as $|f|$ need not be smooth when f is a smooth function—but continuity is good enough for defining the integral $\int_A |\omega|$ as in Theorem 10.30. Changing ω to $|\omega|$ has the effect of replacing f with $|f|$ in the calculation above, and similarly in all other oriented charts. The same argument as above then proves

$$\int_{x(A)} (\varphi^N \circ x^{-1}) |f|\, dm \to \int_A |\omega| \qquad \text{as } N \to \infty.$$

Since φ^N equals 1 on subsets that exhaust all of A as $N \to \infty$, this implies a uniform upper bound for the integral of $|f|$ over arbitrary compact subsets of $x(A)$, and thus $\int_{x(A)} |f| \, dm < \infty$. □

EXERCISE 11.4. For every oriented n-manifold M with $n \geqslant 1$, there exists another oriented manifold $-M$ that is defined as the same manifold with the "reversed" orientation, meaning that one changes the orientation of every tangent space T_pM. Show that for every $\omega \in \Omega_c^n(M)$,

$$\int_{-M} \omega = -\int_M \omega.$$

Hint: If you fix the reflection map $r(t^1, t^2, \ldots, t^n) := (-t^1, t^2, \ldots, t^n)$ *on* $\mathbb{R}^n$ *and take any oriented chart* $(\mathcal{U}, x)$ *on* M*, then* $(\mathcal{U}, r \circ x)$ *will be an oriented chart on* $-M$*.*

REMARK 11.5. At long last, we can now clarify a notational issue that often bothers newcomers to integral calculus: what does $\int_b^a f(x) \, dx$ actually mean when $a < b$? It is traditional to define this as a synonym for $-\int_a^b f(x) \, dx := -\int_{[a,b]} f \, dm$ and regard it as a meaningless but useful convention, but now we can assign a deeper meaning to it: for the 1-manifold $M := (a, b) \subset \mathbb{R}$ with its canonical orientation and the 1-form $f \, dx \in \Omega_c^1(M)$ defined via the canonical coordinate x and a compactly supported[39] function $f : (a, b) \to \mathbb{R}$, the correct definition is

$$\int_b^a f(x) \, dx := \int_{-(a,b)} f \, dx,$$

where $-(a, b)$, denotes the manifold (a, b) with the opposite of its canonical orientation. This is consistent with the way that substitution is typically applied in calculations of 1-dimensional integrals: orientation-reversing diffeomorphisms are sometimes used for substitution, but they produce integrals over intervals with reversed orientation.

11.3. Volume forms. We now consider the first true geometric application of integration: how does one compute volumes of subsets in a manifold?

In an ordinary measure space X with measure μ, the volume of $A \subset X$ is simply $\int_A d\mu$. We have seen that in n-dimensional oriented manifolds, the role of measures is played by differential n-forms; however, not all of these define geometrically appropriate notions of volume. Indeed, a form $\omega \in \Omega^n(M)$ gives a way to define volumes of paralelepipeds in each tangent space T_pM, but it can happen that $\omega_p = 0$ at some point $p \in M$, implying that all regions in T_pM have volume zero, which is not very reasonable geometrically. The objects that we will refer to as "volume forms" specifically exclude this possibility:

DEFINITION 11.6. A **volume form** (*Volumenform*) on an n-manifold M is an n-form $\omega \in \Omega^n(M)$ such that $\omega_p \neq 0$ for all $p \in M$.

NOTATION. In these notes, we will usually denote volume forms by

$$d\mathrm{vol} \in \Omega^n(M),$$

or sometimes $d\mathrm{vol}_M$ if there are various manifolds in the picture and we want to specify which one $d$vol is defined on. The notation is slightly misleading since in many cases, our volume form will not actually be the exterior derivative of anything; nonetheless, the presence of the symbol "d" is consistent with the way that measures are often written in integrals, and that is the role that we intend for $d$vol to play.

[39] We are assuming compact support in (a, b) here because we have not yet defined manifolds *with boundary*, and thus cannot define an integral over the *closed* interval $[a, b]$. This will come in the next lecture, however.

Observe that since $\dim \Lambda^n T_p^* M = 1$ for every $p \in M$, $d\mathrm{vol} := \omega \in \Omega^n(M)$ is a volume form if and only if ω_p is a basis of $\Lambda^n T_p^* M$ for every p, and it follows in this case that any other n-form $\alpha \in \Omega^n(M)$ can be written as

$$\alpha = f\, d\mathrm{vol}$$

for a unique function $f \in C^\infty(M)$. In this situation, α is also a volume form if and only if the function f is nowhere zero.

PROPOSITION 11.7. *Any volume form $d\mathrm{vol} \in \Omega^n(M)$ on a manifold M determines a unique orientation of M such that for each $p \in M$, an ordered basis $(X_1, \ldots, X_n) \in T_p M$ is positively oriented if and only if $d\mathrm{vol}(X_1, \ldots, X_n) > 0$.*

PROOF. Assuming $d\mathrm{vol}_p \neq 0$, Proposition 9.2 implies that $d\mathrm{vol}(X_1, \ldots, X_n) \neq 0$ for every basis $X_1, \ldots, X_n$ of $T_p M$. It follows that $d\mathrm{vol}$ determines a continuous map $\mathcal{B}(T_p M) \to \mathbb{R} : (X_1, \ldots, X_n) \mapsto d\mathrm{vol}(X_1, \ldots, X_n)$ that is never zero, and since it clearly can take values of both signs, it must take positive values on one connected component of $\mathcal{B}(T_p M)$ and negative values on the other. Since its values also vary continuously with the point p, this distinction between the signs of $d\mathrm{vol}(X_1, \ldots, X_n)$ determines a continuous family of orientations of the tangent spaces $T_p M$. $\square$

If M is equipped with the orientation determined by a volume form $d\mathrm{vol}$ via Proposition 11.7, then it is common to write this condition as

$$d\mathrm{vol} > 0,$$

meaning literally that $d\mathrm{vol}(X_1, \ldots, X_n) > 0$ for every $p \in M$ and every *positively-oriented* basis $(X_1, \ldots, X_n)$ of $T_p M$, and $d\mathrm{vol}$ is in this case called a **positive volume form** on the oriented manifold M. Another n-form $\alpha = f\, d\mathrm{vol}$ is then also a positive volume form if and only if $f > 0$ everywhere. In particular, for any oriented chart $(\mathcal{U}, x)$, $dx^1 \wedge \ldots \wedge dx^n$ is a positive volume form on $\mathcal{U}$ since $(dx^1 \wedge \ldots \wedge dx^n)(\partial_1, \ldots, \partial_n) = 1$, thus a positive volume form $d\mathrm{vol} \in \Omega^n(M)$ always locally takes the form

$$d\mathrm{vol} = f\, dx^1 \wedge \ldots \wedge dx^n, \qquad f : \mathcal{U} \to (0, \infty). \tag{11.3}$$

If $(M, d\mathrm{vol})$ is an oriented manifold equipped with a positive volume form, the volume of a measurable subset $A \subset M$ is now defined simply as

$$\mathrm{Vol}(A) := \int_A d\mathrm{vol},$$

which is always nonnegative due to (11.3).

The definition of volume in M clearly depends on a choice of volume form, and for arbitrary manifolds there is generally no canonical choice—this reflects the fact that volumes of regions can appear very different when viewed in different coordinate systems. However, there are situations in which extra data determines a natural choice of volume form.

Suppose for instance that $M \subset \mathbb{R}^n$ is a k-dimensional submanifold of Euclidean space. Each tangent space $T_p M$ is then a k-dimensional linear subspace of $T_p \mathbb{R}^n = \mathbb{R}^n$, and can thus be assigned the standard Euclidean inner product $\langle\ ,\ \rangle$, which we can then use to define lengths of vectors in $T_p M$ and angles between them. In particular, this defines the notion of an *orthonormal* basis of $T_p M$. The paralelepiped spanned by an orthonormal basis of a k-dimensional subspace in $\mathbb{R}^n$ has the same dimensions as the k-dimensional unit cube, so its k-dimensional volume is 1, and it would therefore be natural to choose a volume form $d\mathrm{vol} \in \Omega^k(M)$ that evaluates to 1 on some orthonormal basis.

To bring this discussion into its most natural setting, recall that a **Riemannian metric** (*Riemannsche Metrik*) on a manifold M is a smooth type $(0, 2)$ tensor field $g \in \Gamma(T_2^0 M)$ such that $g_p : T_p M \times T_p M \to \mathbb{R}$ defines an inner product on $T_p M$ for every $p \in M$. The pair (M, g)

is in this case called a **Riemannian manifold** (*Riemannsche Mannigfaltigkeit*). The data of a Riemannian metric makes it possible to define norms of tangent vectors and angles between them, so in particular, every tangent space T_pM acquires a well-defined notion of orthonormality.

DEFINITION 11.8. On a Riemannian manifold (M, g), a volume form $d\mathrm{vol} \in \Omega^n(M)$ is said to be **compatible** with the metric g if for every $p \in M$ and every orthonormal basis $X_1, \ldots, X_n \in T_pM$, $|d\mathrm{vol}(X_1, \ldots, X_n)| = 1$.

Since $\dim \Lambda^n T_p^*M = 1$ for an n-manifold M, there are clearly at most two volume forms compatible with a given metric g at any given point $p \in M$. The following algebraic lemma guarantees that there are, in fact, exactly two, corresponding to the two possible orientations of T_pM.

LEMMA 11.9. *Suppose V is an n-dimensional oriented vector space equipped with an inner product $\langle\ ,\ \rangle$, $v_1, \ldots, v_n \in V$ is a positively-oriented orthonormal basis and $v_*^1, \ldots, v_*^n \in V^*$ denotes its dual basis. Then the top-dimensional form*

$$\omega := v_*^1 \wedge \ldots \wedge v_*^n \in \Lambda^n V^*$$

satisfies $\omega(w_1, \ldots, w_n) = 1$ for every positively-oriented orthonormal basis $w_1, \ldots, w_n \in V$.

PROOF. By (9.3), it will suffice to establish that if $w_*^1, \ldots, w_*^n \in V^*$ is the dual basis of another positively-oriented orthonormal basis $w_1, \ldots, w_n \in V$, then

$$v_*^1 \wedge \ldots \wedge v_*^n = w_*^1 \wedge \ldots \wedge w_*^n.$$

By Proposition 9.10, the scaling factor relating these two n-forms is the determinant of the matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ with entries $A^i{}_j := w_*^i(v_j)$. Writing v_k as a linear combination of the w_i gives $v_k = w_*^i(v_k)w_i$, and orthonormality then implies

$$\begin{aligned}\delta_{k\ell} &= \langle v_k, v_\ell \rangle = \langle w_*^i(v_k)w_i, w_*^j(v_\ell)w_j \rangle = w_*^i(v_k)w_*^j(v_\ell)\langle w_i, w_j \rangle = w_*^i(v_k)w_*^j(v_\ell)\delta_{ij} \\ &= \sum_{i=1}^n w_*^i(v_k)w_*^i(v_\ell) = \sum_{i=1}^n A^i{}_k A^i{}_\ell,\end{aligned}$$

where in the second line we can no longer use the summation convention since the index to be summed does not appear in an upper-lower pair. This calculation implies that the rows of $\mathbf{A}$ form an orthonormal set, meaning $\mathbf{A} \in \mathrm{O}(n)$ and thus $\det(\mathbf{A}) = \pm 1$. Since both bases are also positively oriented, there exists a continuous family of orthonormal bases connecting one to the other, implying that there is also a continuous family of orthogonal matrices connecting $\mathbf{A}$ to $\mathbb{1}$, thus $\det(\mathbf{A}) = 1$. □

COROLLARY 11.10. *Every oriented Riemannian n-manifold (M, g) admits a unique so-called **Riemannian volume form** $d\mathrm{vol} \in \Omega^n(M)$ that is positive and compatible with g.*

PROOF. The existence and uniqueness of $d\mathrm{vol}_p \in \Lambda^n T_p^*M$ for each $p \in M$ follows from Lemma 11.9, so it remains only to check that the n-form defined in this way is smooth. To see this, note that for any $p \in M$, one can find a neighborhood $\mathcal{U} \subset M$ of p and smooth vector fields $X_1, \ldots, X_n \in \mathfrak{X}(\mathcal{U})$ that form a positively-oriented orthonormal basis at every point in $\mathcal{U}$; simply start e.g. with a basis of coordinate vector fields near p and then use the Gram-Schmidt process to make them orthonormal at each point. Now if $\lambda^1, \ldots, \lambda^n \in \Omega^1(\mathcal{U})$ are defined so that $\lambda_q^1, \ldots, \lambda_q^n \in T_q^*M$ is the dual basis to $X_1(q), \ldots, X_n(q) \in T_qM$ for every $q \in \mathcal{U}$, then $\lambda^1 \wedge \ldots \wedge \lambda^n$ is a smooth n-form on $\mathcal{U}$ that matches $d\mathrm{vol}$ according to Lemma 11.9. □

Example 11.11. On $\mathbb{R}^n$, there is a standard choice of Riemannian metric defined by assigning to each $T_p\mathbb{R}^n = \mathbb{R}^n$ the Euclidean inner product. This makes the standard coordinate vector fields $\partial_1, \ldots, \partial_n$ into a positively-oriented orthonormal basis at every point, and the unique positive volume form compatible with the standard metric is thus the so-called **standard volume form** $dx^1 \wedge \ldots \wedge dx^n$. The notion of volume defined by integrating it is of course just the Lebesgue measure.

Exercise 11.12. In local coordinates with respect to an oriented n-dimensional chart $(\mathcal{U}, x)$, a Riemannian metric $g \in \Gamma(T_2^0 M)$ is described in terms of its components $g_{ij} := g(\partial_i, \partial_j)$, so that vectors $X, Y \in T_pM$ at points $p \in \mathcal{U}$ satisfy $g(X,Y) = g_{ij}X^iY^j$. The goal of this exercise is to prove that the Riemannian volume form is then given by

$$\text{dvol} = \sqrt{\det \mathbf{g}}\, dx^1 \wedge \ldots \wedge dx^n \qquad \text{on } \mathcal{U}, \tag{11.4}$$

where $\mathbf{g} : \mathcal{U} \to \mathbb{R}^{n \times n}$ denotes the matrix-valued function whose ith row and jth column is g_{ij}. Note that this matrix necessarily has positive determinant since g is positive definite. Fix a point $p \in \mathcal{U}$ and a positively-oriented orthonormal basis $(X_1, \ldots, X_n)$ of T_pM, whose dual basis we will denote by $\lambda^1, \ldots, \lambda^n \in T_p^*M$. According to Lemma 11.9, $\text{dvol}_p = \lambda^1 \wedge \ldots \wedge \lambda^n$. Define matrices $\mathbf{X}, \boldsymbol{\lambda} \in \mathbb{R}^{n \times n}$ whose ith row and jth column are $dx^i(X_j)$ and $\lambda^i(\partial_j)$ respectively. By Proposition 9.10, $(\lambda^1 \wedge \ldots \wedge \lambda^n)(\partial_1, \ldots, \partial_n) = \det \boldsymbol{\lambda}$.

(1) Prove $\boldsymbol{\lambda} = \mathbf{X}^{-1}$.
(2) Prove $\mathbf{X}^T \mathbf{g} \mathbf{X} = \mathbb{1}$.
(3) Deduce (11.4).

Most people's favorite manifolds are submanifolds of Euclidean space—especially surfaces in $\mathbb{R}^3$. Generalizing this notion slightly, an $(n-1)$-dimensional submanifold M of an n-manifold N is called a **hypersurface** (*Hyperfläche*) in N. Any Riemannian metric g on N induces a Riemannian metric on any submanifold $M \subset N$, defined simply by restricting each of the inner products g_p on tangent spaces T_pN to the subspaces $T_pM \subset T_pN$. To put this another way, one can denote the inclusion map of M into N by $i : M \hookrightarrow N$ and observe that for every $p \in M$, $i_* : T_pM \hookrightarrow T_pN$ is the corresponding inclusion map of vector spaces, so the Riemannian metric induced by $g \in \Gamma(T_2^0 N)$ on M is the pullback $i^*g \in \Gamma(T_2^0 M)$. With this understood, we will show next that there is an easy way to derive from the compatible volume form on an oriented Riemannian manifold the compatible volume form on any oriented hypersurface.

Definition 11.13. For an n-dimensional vector space V and an integer $k = 1, \ldots, n$, the **interior product** is the bilinear map

$$V \times \Lambda^k V^* \to \Lambda^{k-1} V^* : (v, \alpha) \mapsto \iota_v \alpha$$

defined by $\iota_v\alpha(w_1, \ldots, w_{k-1}) := \alpha(v, w_1, \ldots, w_{k-1})$. On a manifold M, the map

$$\mathfrak{X}(M) \times \Omega^k(M) \to \Omega^{k-1}(M) : (X, \omega) \mapsto \iota_X \omega$$

is defined similarly by $(\iota_X\omega)_p := \iota_{X(p)}\omega_p$ for all $p \in M$.

Proposition 11.14. *Assume (N, g) is a Riemannian manifold, $M \subset N$ is a hypersurface with inclusion map $i : M \hookrightarrow N$, and $\nu : M \to TN$ is a continuous map*[40] *such that for every $p \in M$, $\nu(p) \in T_pN$ is a unit vector orthogonal to T_pM. (In this situation we call ν a* ***unit normal vector field*** *for M.) Then if $\text{dvol}_N \in \Omega^n(N)$ is a volume form on N compatible with g,*

$$\text{dvol}_M := (\iota_\nu \text{dvol}_N)|_{TM} \in \Omega^{n-1}(M)$$

[40] In fact it will follow from these assumptions that ν is also smooth, but one does not need to know that in advance.

is a volume form on M compatible with the induced metric i^*g.

PROOF. For any $p \in M$ and an orthonormal basis $X_1, \ldots, X_{n-1}$ of T_pM, the n-tuple $\nu(p), X_1, \ldots, X_{n-1}$ forms an orthonormal basis of T_pN, thus

$$|\iota_\nu \mathrm{dvol}_N(X_1, \ldots, X_{n-1})| = |\mathrm{dvol}(\nu(p), X_1, \ldots, X_{n-1})| = 1.$$

□

EXERCISE 11.15. Using Cartesian coordinates (x, y, z) on $\mathbb{R}^3$, let $\omega := x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy \in \Omega^2(\mathbb{R}^3)$, and let $i : S^2 \hookrightarrow \mathbb{R}^3$ denote the inclusion of the unit sphere.

(a) Show that $d\mathrm{vol}_{S^2} := i^*\omega \in \Omega^2(S^2)$ is a volume form compatible with the Riemannian metric on S^2 induced by the Euclidean inner product.
Hint: Pick a good vector field $X \in \mathfrak{X}(\mathbb{R}^3)$ with which to write ω as $\iota_X(dx \wedge dy \wedge dz)$.
(b) Show that in the spherical coordinates (θ, ϕ) of Exercise 1.7, $d\mathrm{vol}_{S^2} = \cos\phi\, d\theta \wedge d\phi$.
(c) On the open upper hemisphere $\mathcal{U}_+ := \{z > 0\} \subset S^2 \subset \mathbb{R}^3$, one can define a chart $(x, y) : \mathcal{U}_+ \to \mathbb{R}^2$ by restricting to $\mathcal{U}_+$ the usual Cartesian coordinates x and y, which are then related to the z-coordinate on this set by $z = \sqrt{1 - x^2 - y^2}$. Show that $d\mathrm{vol}_{S^2} = \frac{1}{z}\, dx \wedge dy$ on $\mathcal{U}_+$.
(d) Compute the surface area of $S^2 \subset \mathbb{R}^3$ in two ways: once using the formula for $d\mathrm{vol}_{S^2}$ in part (b), and once using part (c) instead. In both cases, the results of §11.2 will allow you to express the answer in terms of a single Lebesgue integral over a region in $\mathbb{R}^2$, and there will be no need for any partition of unity.

11.4. Densities. [41]

You may have wondered: what if M is non-orientable, but I still want to compute its volume?

There are two problems in this situation: one is that according to Proposition 11.7, M cannot admit a volume form if it does not also admit an orientation, but there is also the more fundamental issue that the integral of an n-form over an n-manifold is not defined unless M comes with an orientation. Recall from §10.1: the trouble was that if $\omega = f\,dx^1 \wedge \ldots \wedge dx^n = g\,dy^1 \wedge \ldots \wedge dy^n$ for two different local coordinate systems $x, y : \mathcal{U} \to \mathbb{R}^n$ on the same region, then the Legesgue integrals $\int_{x(\mathcal{U}\cap A)} f \circ x^{-1}\, dm$ and $\int_{y(\mathcal{U}\cap A)} g \circ y^{-1}\, dm$ cannot generally be assumed to match unless the transition map $\psi := y \circ x^{-1} : x(\mathcal{U}) \to y(\mathcal{U})$ is orientation preserving. This problem is summarized by Equation (10.3), which resembles the classical change-of-variables formula, but does not match it exactly unless $\det(D\psi)$ is everywhere positive.

One way to circumvent this problem is to give up on integrating the real-valued functions f and g and instead integrate their absolute values, so that (10.3) gives rise to the completely true statement

$$\int_{y(A)} \left|g \circ y^{-1}\right| dm = \int_{\psi(x(A))} |G|\, dm = \int_{x(A)} |(G \circ \psi)| \cdot |\det D\psi|\, dm = \int_{x(A)} \left|f \circ x^{-1}\right| dm,$$

in which we are again writing $G := g \circ y^{-1}$. The message of this calculation is that if we are willing to ignore the sign of an n-form and pay attention only to its magnitude, then we will no longer need to restrict ourselves to orientation-preserving transition maps.

DEFINITION 11.16. A (nonnegative) **density** on a smooth n-manifold M is a map

$$\mu : (TM)^{\oplus n} \to [0, \infty)$$

[41] The contents of §11.4 were not covered in the lecture and will not be referred to again in this course, at least not in any serious way. This section of the notes is provided only for your information.

whose restriction to $T_pM \times \ldots \times T_pM$ for each $p \in M$ takes the form $\mu_p(X_1, \ldots, X_n) = |\omega_p(X_1, \ldots, X_n)|$ for some $\omega_p \in \Lambda^n T_p^*M$. In a smooth chart $(\mathcal{U}, x)$, every density can thus be written in terms of the standard volume form $dx^1 \wedge \ldots \wedge dx^n \in \Omega^n(\mathcal{U})$ as

$$\mu = f \cdot \left|dx^1 \wedge \ldots \wedge dx^n\right|$$

for a unique function $f : \mathcal{U} \to [0, \infty)$. We call μ a **smooth density** if the function f defined in this way is smooth for all choices of smooth chart on M.

REMARK 11.17. It is also possible to define densities with negative values (see e.g. [Lee13]), but we will not need this. Our refusal to define negative densities means that the space

$$\mathscr{D}(M) := \{\text{smooth densities on } M\}$$

is not a vector space, but it does admit natural notions of addition and multiplication by nonnegative scalars.

The **support** of a density $\mu \in \mathscr{D}(M)$ is of course the closure of the set $\{p \in M \mid \mu_p \neq 0\} \subset M$, and we will denote

$$\mathscr{D}_c(M) := \left\{\mu \in \mathscr{D}(M) \;\middle|\; \mu \text{ has compact support}\right\}.$$

For smooth maps $\varphi : M \to N$, there is a natural **pullback** operation $\varphi^* : \mathscr{D}(N) \to \mathscr{D}(M)$ defined by

$$(\varphi^*\mu)(X_1, \ldots, X_n) := \mu(\varphi_*X_1, \ldots, \varphi_*X_n).$$

If we revise the discussion of §10.1 to work with densities instead of n-forms, then the key fact is that for any two charts x and y defined on the same domain $\mathcal{U}$, we have

$$\left|dy^1 \wedge \ldots \wedge dy^n\right| = \left|\det\left(\frac{\partial y}{\partial x}\right)\right| \cdot \left|dx^1 \wedge \ldots \wedge dx^n\right| \qquad \text{on } \mathcal{U},$$

thus if $\mu = f\,|dx^1 \wedge \ldots \wedge dx^n| = g\,|dy^1 \wedge \ldots \wedge dy^n|$ on this region, the nonnegative functions f and g are related by $f = g \cdot \left|\det\left(\frac{\partial y}{\partial x}\right)\right|$. The presence of the absolute value in this expression repairs our previous problem with orientations, and it now follows that the integrals $\int_{x(A)} f \circ x^{-1}\,dm$ and $\int_{y(A)} g \circ y^{-1}\,dm$ will *always* match, even if $y \circ x^{-1}$ is orientation reversing. The proof of Theorem 10.30 can now easily be adapted to establish the following:

THEOREM 11.18. *For $n \in \mathbb{N}$, one can uniquely associate to every smooth n-manifold M and measurable subset $A \subset M$ a map*

$$\mathscr{D}_c(M) \to [0, \infty) : \mu \mapsto \int_A \mu$$

such that the following conditions are satisfied:

(1) $\int_A (\mu_1 + \mu_2) = \int_A \mu_1 + \int_A \mu_2$ for any $\mu_1, \mu_2 \in \mathscr{D}_c(M)$.
(2) If $\mathcal{U} \subset M$ is an open subset containing $\operatorname{supp}(\mu) \cap A$, then then $\int_{\mathcal{U} \cap A} \mu = \int_A \mu$.
(3) For $M = \mathcal{U} \subset \mathbb{R}^n$ an open subset of Euclidean space and the standard Cartesian coordinates $x^1, \ldots, x^n$,

$$\int_A f\,\left|dx^1 \wedge \ldots \wedge dx^n\right| = \int_A f\,dm$$

for all smooth compactly supported functions $f : \mathcal{U} \to [0, \infty)$, where the right hand side is the standard Lebesgue integral of f.

(4) For any diffeomorphism $\psi : M \to N$ between a pair of n-manifolds,

$$\int_A \psi^* \mu = \int_{\psi(A)} \mu$$

holds for all $\mu \in \mathscr{D}_c(N)$ and measurable subsets $A \subset M$.

□

The freedom in this theorem to allow non-orientable manifolds and diffeomorphisms that are not orientation preserving is paid for by the fact that integrals of nonnegative densities are always nonnegative, and thus tend to deliver less information than the *real*-valued integrals of differential forms. As mentioned in Remark 11.17 above, one can also allow densities with negative values and thus obtain negative integrals, but this does not add very much generality: it is tantamount to defining a measure μ via integrals of a positive density and then computing integrals $\int_A f\, d\mu$ of functions f that are also allowed to have negative values. Integration of densities is a somewhat less elegant and less useful construction on the whole than integration of forms; in particular, there are many more beautiful theorems involving the latter. Nonetheless, there are of course geometric situations in which an integral that is guaranteed to be nonnegative is exactly what one wants:

Definition 11.19. A **volume element** on a smooth n-manifold M is a density $d\mathrm{vol}$ such that $d\mathrm{vol}_p \neq 0$ for every $p \in M$. If M is equipped with a volume element $d\mathrm{vol}$, one defines the **volume** of measurable sets $A \subset M$ by

$$\mathrm{Vol}(A) := \int_A d\mathrm{vol} \geqslant 0.$$

We can now state a version of Corollary 11.10 that does not depend on orientability; its proof is an easy adaptation of arguments in the previous section.

Proposition 11.20. *Every Riemannian manifold (M, g) admits a unique volume element $d\mathrm{vol}$ such that for all $p \in M$ and every orthonormal basis $X_1, \ldots, X_n$ of T_pM, $d\mathrm{vol}(X_1, \ldots, X_n) = 1$.* □

We will not have any more occasions to talk about densities and volume elements in this course, but it is good to be aware that a theory of integration exists for non-orientable manifolds, even if it is less versatile and less powerful than the orientable case.

12. Stokes' theorem

It is finally time to tell you the true reason why the exterior derivative is important: it is "dual" in some sense to the operation of replacing a manifold by its boundary. First we will have to discuss what is meant by the *boundary* of a manifold, and we will have to be fairly careful with orientations if we want to get all the signs right.

12.1. A word about dimension zero. You may or may not have noticed that manifolds of dimension zero have been explicitly excluded from all discussion of orientations and integration so far. You probably didn't miss it, because in truth, integrals of 0-forms on 0-manifolds are not very interesting. But we have to define them now, because as soon as we start talking about manifolds with boundary, 0-manifolds will inevitably arise, namely as boundaries of 1-manifolds.

A 0-manifold M, you may recall, is simply a discrete set, and it can have at most countably many elements; it is compact if and only if it is finite. A 0-form on M is then an arbitrary function $f : M \to \mathbb{R}$. There is no need to worry about continuity or smoothness since M is discrete, and the support of f is just the set of all points p where $f(p) \neq 0$, so $f : M \to \mathbb{R}$ has compact support if and only if it is zero outside of a finite set.

Since there is no such thing as a "basis" of a 0-dimensional vector space and no meaningful sense in which one can say that a (the) map $\mathbb{R}^0 \to \mathbb{R}^0$ preserves or reverses orientation, the entire discussion of orientations in §10.2 is useless for $n = 0$. What we will use instead looks terribly naive at first glance, but we will see that it works:

DEFINITION 12.1. An **orientation** of a 0-manifold M is a function $\varepsilon : M \to \{1, -1\}$, i.e. it a assigns to each point of M a label as either "positive" or "negative". A bijection $\varphi : M \to N$ between two oriented 0-manifolds is **orientation preserving** if it maps all positive points to positive points and all negative points to negative points, and it is **orientation reversing** if it exchanges the sets of positive and negative points.

DEFINITION 12.2. For M a 0-manifold with orientation $\varepsilon : M \to \{1, -1\}$ and $f \in \Omega^0_c(M)$, the integral of f on a subset $A \subset M$ is defined by

$$\int_A f := \sum_{p \in A} \varepsilon(p) f(p),$$

where the sum is necessarily finite since f has compact support.

The only other thing worth saying for now about this definition is that it trivially satisfies the usual change-of-variables formula

$$\int_A \varphi^* f = \int_{\varphi(A)} f, \qquad f \in \Omega^0_c(N)$$

whenever $\varphi : M \to N$ is an orientation-preserving bijection of oriented 0-manifolds.

12.2. Manifolds with boundary. The definitions from Lectures 1 and 2 need to be generalized if we want to accommodate examples like the unit n**-disk**

$$\mathbb{D}^n := \left\{ x \in \mathbb{R}^n \;\middle|\; |x| \leqslant 1 \right\},$$

whose interior is accurately described as a smooth n-manifold, but there are no n-dimensional charts (by our current definition) describing neighborhoods in $\mathbb{D}^n$ of points on the boundary

$$\partial \mathbb{D}^n := S^{n-1} \subset \mathbb{D}^n.$$

An even simpler example is the **half-plane**

$$\mathbb{H}^n := (-\infty, 0] \times \mathbb{R}^{n-1} \subset \mathbb{R}^n,$$

whose boundary is the linear subspace

$$\partial \mathbb{H}^n := \{0\} \times \mathbb{R}^{n-1} \subset \mathbb{R}^n.$$

Just as subspaces of this form serve as local models of submanifolds as seen through slice charts, the half-plane will serve as our local model for a manifold with boundary.

DEFINITION 12.3. An n-dimensional **boundary chart** $(\mathcal{U}, x)$ on a set M consists of a subset $\mathcal{U} \subset M$ and an injective map $x : \mathcal{U} \hookrightarrow \mathbb{H}^n$ whose image $x(\mathcal{U}) \subset \mathbb{H}^n$ is an open set.[42]

The only difference between this and Definition 1.4 is the replacement of $\mathbb{R}^n$ by the half-space $\mathbb{H}^n$. A boundary chart $(\mathcal{U}, x)$ will sometimes also be a chart according to our original definition, because an open subset $x(\mathcal{U}) \subset \mathbb{H}^n$ might also be an open subset of $\mathbb{R}^n$; indeed, it will be so if $x(\mathcal{U}) \cap \partial \mathbb{H}^n = \varnothing$. For this reason, any set that is covered by charts can equally well be covered by boundary charts: one need only modify each chart $(\mathcal{U}, x)$ by a translation so that its

[42]One finds a few variations on this definition in the literature, in which the half-space $\mathbb{H}^n = (-\infty, 0] \times \mathbb{R}^{n-1}$ gets replaced by different half-spaces such as $[0, \infty) \times \mathbb{R}^{n-1}$ or $\mathbb{R}^{n-1} \times [0, \infty)$. This detail makes no meaningful difference for the definition of a smooth manifold with boundary, but it starts to matter as soon as one has to think about orientations. The definition in the form we've given here leads to the simplest possible definition of boundary orientations, and a relatively straightforward proof of Stokes' theorem.

image lies in the interior of the half-plane, or if this is impossible because $x(\mathcal{U})$ is unbounded in the x^1-direction, first break it up into countably many open subsets so that this can be done. However, if $x(\mathcal{U})$ does contain points in the boundary $\partial\mathbb{H}^n$, then it is not open in $\mathbb{R}^n$. A typical example is the "open" half-disk

$$\mathring{\mathbb{D}}^n_- := \left\{(x^1, \dots, x^n) \in \mathbb{R}^n \;\middle|\; (x^1)^2 + \dots + (x^n)^2 < 1 \text{ and } x^1 \leqslant 0\right\},$$

which is open in $\mathbb{H}^n$ but not open in $\mathbb{R}^n$ since it does not contain any ball around points in $\mathring{\mathbb{D}}^n_- \cap \partial\mathbb{H}^n$. In this sense, Definition 12.3 is strictly more general than our original definition of a chart.

The notion of **transition maps** between two charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ generalizes in an obvious way to boundary charts,

$$\begin{aligned} \mathbb{H}^n \supset x(\mathcal{U} \cap \mathcal{V}) &\xrightarrow{y \circ x^{-1}} y(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{H}^n, \\ \mathbb{H}^n \supset y(\mathcal{U} \cap \mathcal{V}) &\xrightarrow{x \circ y^{-1}} x(\mathcal{U} \cap \mathcal{V}) \subset \mathbb{H}^n, \end{aligned} \tag{12.1}$$

though since $x(\mathcal{U} \cap \mathcal{V})$ and $y(\mathcal{U} \cap \mathcal{V})$ may be open in $\mathbb{H}^n$ but not in $\mathbb{R}^n$, the notion of smooth compatibility requires a bit of clarification. The quickest approach is to say that a map $f : \mathcal{O} \to \mathbb{R}^m$ defined on some (not necessarily open) subset $\mathcal{O} \subset \mathbb{R}^n$ is of class C^k if and only if it admits an extension of class C^k to some open neighborhood of $\mathcal{O}$ in $\mathbb{R}^n$. With this understood, we will call $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ **smoothly compatible** if both of the transition maps in (12.1) admit smooth extensions over open (in $\mathbb{R}^n$) neighborhoods of their domains.

Remark 12.4. For open subsets $\mathcal{O} \subset \mathbb{H}^n$ in half-space, the notion of a C^k-map $f : \mathcal{O} \to \mathbb{R}^m$ admits various alternative characterizations that do not require extending f over a larger neighborhood in $\mathbb{R}^n$. Denote $\partial\mathcal{O} := \mathcal{O} \cap \partial\mathbb{H}^n$ and $\mathring{\mathcal{O}} := \mathcal{O} \backslash \partial\mathcal{O}$. Then $f : \mathcal{O} \to \mathbb{R}^m$ is of class C^k if and only if its restriction $f|_{\mathring{\mathcal{O}}} : \mathring{\mathcal{O}} \to \mathbb{R}^m$ is of class C^k and either of the following equivalent conditions are satisfied:

- All partial derivatives of $f|_{\mathring{\mathcal{O}}} : \mathring{\mathcal{O}} \to \mathbb{R}^m$ up to order k admit continuous extensions over $\mathcal{O}$;
- All partial derivatives of $f|_{\mathring{\mathcal{O}}} : \mathring{\mathcal{O}} \to \mathbb{R}^m$ up to order k are uniformly continuous on bounded subsets of $\mathring{\mathcal{O}}$.

It is an easy analysis exercise to show that these two conditions are equivalent, and they clearly also follow from the assumption that $f : \mathcal{O} \to \mathbb{R}^m$ admits a C^k-extension to a neighborhood, but the converse takes more effort to prove. We will not do so here since we will never need to use this fact, but the details can be found e.g. in [**AF03**, §5.19–§5.21].

A **smooth n-manifold with boundary** can now be defined by generalizing our previous definition of a smooth n-manifold so that all charts in its maximal smooth atlas are allowed to be boundary charts. Implicit in this definition is the fact that an atlas of boundary charts on M determines a natural topology on M such that the domains of boundary charts are also open sets in M and the charts themselves are homeomorphisms onto their images. This definition is *strictly* more general than what we have been working with so far: a manifold with boundary can sometimes also be a manifold in our previous sense, because its atlas might consist only of regular charts whose images are open subsets of $\mathbb{R}^n$. But if M is a manifold with boundary, it contains a distinuished subset

$$\partial M := \left\{p \in M \;\middle|\; x(p) \in \partial\mathbb{H}^n \text{ for some smooth boundary chart } (\mathcal{U}, x)\right\},$$

called its **boundary** (*Rand*). It should be easy to convince yourself that if $x(p) \in \partial\mathbb{H}^n$ for some particular boundary chart $(\mathcal{U}, x)$, then this also holds for every other boundary chart $(\mathcal{V}, y)$ with $p \in \mathcal{V}$; this is because by the inverse function theorem, the transition maps in (12.1) necessarily preserve the interior of $\mathbb{H}^n$, and therefore also preserve its boundary $\partial\mathbb{H}^n$. Moreover, every boundary chart whose domain intersects ∂M can be viewed as a slice chart for ∂M, so that it is appropriate to call ∂M a smooth $(n-1)$-dimensional submanifold of M. In particular, ∂M inherits from M a natural smooth structure and becomes a smooth $(n-1)$-manifold. We observe that M itself is a manifold in our previous sense if and only if $\partial M = \emptyset$; one sometimes says in this case that M is a *manifold without boundary.* Since $x(\mathcal{U}) \cap \partial\mathbb{H}^n$ is *always* an open subset of $\partial\mathbb{H}^n = \{0\} \times \mathbb{R}^{n-1}$ for a boundary chart $(\mathcal{U}, x)$, the manifold ∂M never has boundary, i.e.

$$\partial(\partial M) = \emptyset.$$

Remark 12.5. One can define even more general notions such as a "manifold with boundary and corners," in which images of charts are allowed to be open subsets of quadrants like $(-\infty, 0] \times (-\infty, 0] \times \mathbb{R}^{n-2}$, in which case ∂M may also be a manifold with nonempty boundary (and possibly corners). The literature on these objects seems however to be not entirely unanimous on what the correct definitions are. In this course, we will occasionally mention corners in heuristic discussions, but we will not study them in any serious way.

Remark 12.6. From now on, you must pay careful attention whenever you see the word "manifold" without further modifiers, as its default meaning may be either "manifold without boundary" or "manifold with boundary" depending on the context. Keep in mind also that these categories are not mutually exclusive: a "manifold with boundary" may have $\partial M = \emptyset$. I generally make a point of saying "manifold with nonempty boundary" if I want to explicitly assume $\partial M \neq \emptyset$. I also will often refer to boundary charts simply as "charts" when working in the context of manifolds with boundary.

Example 12.7. Suppose N is an n-manifold without boundary and $M \subset N$ is an open subset such that $\overline{M}\backslash M \subset N$ is a smooth $(n-1)$-dimensional submanifold, i.e. a hypersurface. Then the closure $\overline{M} \subset N$ is naturally a smooth n-manifold with boundary and

$$\partial\overline{M} = \overline{M}\backslash M,$$

because every slice chart for $\overline{M}\backslash M$ can be modified in straightforward ways so as to be interpreted as a boundary chart for $\overline{M}$. Most interesting examples of manifolds with boundary arise in this way, and it can be shown that *all* manifolds with boundary are diffeomorphic to examples of this type, though the ambient manifold N might not always be a natural part of the picture. As an important special case, if $f : N \to \mathbb{R}$ is a smooth function with $c \in \mathbb{R}$ as a regular value, then $f^{-1}((-\infty, c])$ and $f^{-1}([c, \infty))$ are naturally manifolds with boundary, the boundary in each case being the regular level set $f^{-1}(c) \subset N$. Examples of this type include the n-disk $\mathbb{D}^n \subset \mathbb{R}^n$ mentioned at the beginning of this section.

Almost all of the notions we have discussed in this course so far—tangent vectors and tangent maps, vector fields, tensors, forms, orientations—can be generalized in straightforward ways for manifolds with boundary so long as one remembers what smoothness means on open sets in half-space. The tangent spaces T_pM are defined exactly as before for $p \in M\backslash\partial M$, though it takes a bit more thought to arrive at the right definition for $p \in \partial M$. Here it is useful to keep Example 12.7 in mind and imagine M as a closed subset of a larger manifold N without boundary such that $\partial M \subset N$ is a smooth hypersurface: the correct definition for $p \in \partial M$ is then $T_pM := T_pN$, so that T_pM is still a vector space of the same dimension as M. If there is no ambient manifold N in the picture, then one can instead modify the original definition of T_pM in terms of paths through p by allowing paths of the form $\gamma : (-\epsilon, 0] \to M$ or $\gamma : [0, \epsilon) \to M$ that run "out of" or "into" M through its boundary. The crucial thing to remember is that for any chart $(\mathcal{U}, x)$ with $p \in \mathcal{U}$, $d_px : T_pM \to \mathbb{R}^n$ is still a linear isomorphism, even if $p \in \partial M$. Since $\partial M \subset M$ is an $(n-1)$-dimensional submanifold, $T_p(\partial M) \subset T_pM$ is an $(n-1)$-dimensional subspace. The complement

$T_pM\backslash T_p(\partial M)$ has two connected components: one consists of all vectors that point **outward**, meaning they are derivatives of "departing" paths $\gamma : (-\epsilon, 0] \to M$, and the other contains vectors that point **inward**, which are derivatives of "entering" paths $\gamma : [0, \epsilon) \to M$. It should go without saying that flows of vector fields $X \in \mathfrak{X}(M)$ require extra care when $\partial M \neq \varnothing$, because e.g. if $p \in \partial M$ and $X(p)$ points outward/inward, then there is no forward/backward flow line starting at p for any nonzero time. There is no problem however if $X|_{\partial M}$ is everywhere tangent to the boundary, since it then also defines a flow on ∂M, and Theorem 5.1 in this case goes through without changes.

The notion of a submanifold also requires slight modification when boundaries are involved: the appropriate definition is to call $M \subset N$ a **submanifold** (with boundary) whenever it is the image of an embedding of some manifold with boundary. This allows a few possibilities that were not covered by our original definition in terms of slice charts: one of them was already mentioned above, namely the natural embedding of the boundary $\partial M \hookrightarrow M$. Another is Example 12.7: if N is an n-manifold and $M \subset N$ is an open subset such that $\partial \overline{M} := \overline{M}\backslash M$ is a smooth hypersurface in M, then $\overline{M}$ is a smooth n-dimensional submanifold with boundary in N. This opens the previously excluded possibility that a manifold and submanifold may have the same dimension without one being an open subset of the other.

Proposition 12.8. *If M is an oriented manifold of dimension $n \geqslant 2$ with boundary, then the $(n-1)$-manifold ∂M inherits a natural orientation such that for every oriented boundary chart $(\mathcal{U}, x)$ on M, $(\mathcal{U} \cap \partial M, x|_{\mathcal{U}\cap\partial M})$ is an oriented chart on ∂M. This orientation can also be characterized as follows: for every point $p \in \partial M$ and any tangent vector $\nu \in T_pM\backslash T_p(\partial M)$ that points outward, a basis $(X_1, \ldots, X_{n-1})$ of $T_p(\partial M)$ is positively oriented if and only if the basis $(\nu, X_1, \ldots, X_{n-1})$ of T_pM is positively oriented.*

The orientation defined on ∂M from an orientation of M via this proposition is called the **boundary orientation**. We will always assume unless otherwise specified that when M is oriented, ∂M is endowed with the boundary orientation.

Proof of Proposition 12.8. The main point is that any orientation-preserving transition map $\psi := y \circ x^{-1} : x(\mathcal{U} \cap \mathcal{V}) \to y(\mathcal{U} \cap \mathcal{V})$ not only preserves the subset $\partial\mathbb{H}$ but is also orientation preserving on this subset. To see this, observe that the derivative $D\psi(q) : \mathbb{R}^n \to \mathbb{R}^n$ at any point q must be an isomorphism that preserves each of the subsets $\mathbb{H}^n$ and $\partial\mathbb{H}^n$, thus it is represented by a matrix of the form

$$D\psi(q) = \begin{pmatrix} a & 0 \\ \mathbf{v} & \mathbf{B} \end{pmatrix}, \qquad a > 0,\ \mathbf{v} \in \mathbb{R}^{n-1},\ \mathbf{B} \in \mathbb{R}^{(n-1)\times(n-1)},$$

where $\mathbf{B}$ is the derivative at q of the restricted transition map on $\partial\mathbb{H}$. Clearly $\det D\psi(q) > 0$ if and only if $\det \mathbf{B} > 0$. This shows that the restriction of an oriented atlas of M to ∂M is an oriented atlas of ∂M.

To characterize the boundary orientation in terms of bases, choose any oriented chart $(\mathcal{U}, x)$ near a point $p \in \partial M$, so the coordinate vector fields $\partial_1, \ldots, \partial_n$ define a positively-oriented basis of T_pM. The restriction of $(\mathcal{U}, x)$ to ∂M now defines an oriented chart for ∂M near p, and the coordinate vector fields for this restricted chart are $(\partial_2, \ldots, \partial_n)$, which therefore form a positively-oriented basis of $T_p(\partial M)$, and this can then be deformed continuously through bases to any other positively-oriented basis $(X_1, \ldots, X_{n-1})$ of $T_p(\partial M)$. Since ∂_1 points outward at p, it follows that for any other vector $\nu \in T_pM\backslash T_p(\partial M)$ pointing outward, the basis $(\nu, X_1, \ldots, X_{n-1})$ of T_pM can be deformed continuously through bases to $(\partial_1, \ldots, \partial_n)$, simply by deforming $(X_1, \ldots, X_{n-1})$ through bases of $T_p(\partial M)$ to $(\partial_2, \ldots, \partial_n)$ and simultaneously deforming ν through outward-pointing vectors to ∂_1. This proves that $(\nu, X_1, \ldots, X_{n-1})$ is a positively-oriented basis of T_pM, and conversely, if

$(X_1, \ldots, X_{n-1})$ had been negatively oriented, we could apply the same argument to the positively-oriented basis $(-X_1, X_2, \ldots, X_{n-1})$ and thus conclude that $(\nu, X_1, \ldots, X_{n-1})$ is also negatively oriented. $\square$

We had to exclude the case $\dim M = 1$ from Proposition 12.8 because orientations of 0-manifolds cannot be described in terms of charts or bases.

DEFINITION 12.9. If M is an oriented 1-manifold with boundary, the **boundary orientation** of the 0-manifold ∂M is defined by calling a point $p \in \partial M$ positive if the basis of T_pM formed by an outward-pointing vector $\nu \in T_pM$ is positively oriented, and negative otherwise.

EXAMPLE 12.10. Any nontrivial compact interval $[a, b] \subset \mathbb{R}$ is a 1-manifold with boundary, and if we assign it the canonical orientation of $\mathbb{R}$ then the boundary orientation of $\partial[a, b] = \{a, b\}$ makes b a positive point and a a negative point. Informally, we write

$$\partial[a, b] = -\{a\} \amalg \{b\}.$$

A slightly different example is

$$\partial(-\infty, 0] = \{0\},$$

in which the point 0 is assigned a positive orientation; this will be relevant in the proof of Stokes' theorem below.

12.3. The boundary operator is a graded derivation. I want to point out something about boundary orientations that is not an essential part of this discussion, but it may help you to understand more intuitively why graded Leibniz rules keep showing up.

In the previous section we defined an operator "∂" that takes an oriented n-manifold M (with boundary) and returns an oriented $(n-1)$-manifold ∂M. It satisfies $\partial(\partial M) = \varnothing$ for all M, which seems formally similar to the relation $d \circ d = 0$ satisfied by the exterior derivative. We will see in the next section that the operators ∂ and d are in fact dual to each other in a sense that can be made precise, thus it should not be surprising that they have formally similar properties. We claim in particular that ∂ also satisfies a graded Leibniz rule.

To understand what this means, suppose M and N are two oriented manifolds with boundary, with $\dim M = m$ and $\dim N = n$. This discussion will be heuristic, so we will choose not to worry about the fact that $M \times N$ might not actually be a smooth manifold with boundary: in particular, the neighborhood of a point $(p, q) \in \partial M \times \partial N \subset M \times N$ cannot be described smoothly via our usual notion of a boundary chart, and a completely correct description would require the notion of manifolds with boundary *and corners* (cf. Remark 12.5). Nonetheless, it seems sensible to write

$$\partial(M \times N) = (\partial M \times N) \cup (M \times \partial N)\,, \tag{12.2}$$

and outside of the exceptional subset $\partial M \times \partial N$, it is literally true that $M \times N$ is a smooth manifold whose boundary is the union of these two pieces. Formally, $M \times N$ is a smooth manifold with boundary and corners, and its boundary consists of two smooth *faces* $\partial M \times N$ and $M \times \partial N$, each of which are smooth manifolds with boundary, and they are attached to each other at their common boundary $\partial M \times \partial N$.

Now, let's say all that again but pay attention to orientations. The product of two oriented manifolds M and N carries a natural **product orientation** such that for any $(p, q) \in M \times N$ and any pair of positively oriented bases $(X_1, \ldots, X_m)$ of T_pM and $(Y_1, \ldots, Y_n)$ of T_qN, $(X_1, \ldots, X_m, Y_1, \ldots, Y_n)$ is a positively-oriented basis of $T_{(p,q)}(M \times N) = T_pM \times T_qN$; here we identify each $X_i \in T_pM$ with $(X_i, 0) \in T_pM \times T_qN = T_{(p,q)}(M \times N)$ and similarly identify $Y_j \in T_qN$ with $(0, Y_j) \in T_pM \times T_qN = T_{(p,q)}(M \times N)$. Now, if ∂M and ∂N are each endowed with their natural boundary orientations, then the two faces $\partial M \times N$ and $M \times \partial N$ of the boundary of $M \times N$ inherit product orientations, but these may or may not match the boundary orientation of

$\partial(M \times N)$. Indeed, at a point $(p,q) \in \partial M \times N$, if we choose a positively-oriented basis $(X_2, \ldots, X_m)$ of $T_p(\partial M)$ and an outward-pointing vector $\nu \in T_pM \backslash T_p(\partial M)$, then $(\nu, 0) \in T_{(p,q)}(M \times N)$ also points outward through $\partial M \times N$ and $(\nu, X_2, \ldots, X_m, Y_1, \ldots, Y_n)$ forms a positively-oriented basis of $T_{(p,q)}(M \times N)$, implying that the boundary orientation of $\partial(M \times N)$ does match the product orientation of $\partial M \times N$. But things are different at a point $(p,q) \in M \times \partial N$. Choosing a positively-oriented basis $(Y_2, \ldots, Y_n)$ of $T_q(\partial N)$ and an outward-pointing vector $\nu \in T_qY \backslash T_q(\partial Y)$, a positively-oriented basis of $M \times N$ is given by $(X_1, \ldots, X_m, \nu, Y_2, \ldots, Y_n)$, but m flips are required in order to permute this basis to $(\nu, X_1, \ldots, X_m, Y_2, \ldots, Y_n)$, in which ν serves as an outward-pointing vector in $T_{(p,q)}(M \times N) \backslash T_{(p,q)}(\partial(M \times N))$ and $(X_1, \ldots, X_m, Y_2, \ldots, Y_n)$ as a positively-oriented basis for the product orientation on $M \times \partial N$. This means that the product orientation of $M \times \partial N$ matches the boundary orientation of $\partial(M \times N)$ if and only if $(-1)^m = 1$, i.e. if m is even. The oriented version of (12.2) can thus be written as

$$\partial(M \times N) = (\partial M \times N) \cup ((-1)^m (M \times \partial N)), \tag{12.3}$$

where we define $-(M \times \partial N)$ to mean the oriented manifold obtained from $M \times \partial N$ by assigning it the *opposite* of the product orientation. The formal resemblance of this formula to a graded Leibniz rule is difficult to ignore, though we cannot make this notion precise in the present context since we have not defined any algebraic structure on the "set" of manifolds with boundary and corners. The easiest way to make such notions precise is probably by defining homology theory, which is a topic for a topology course and not for this one, but I wanted in any case to provide (12.3) as further evidence of a formal similarity between the operators ∂ and d.

12.4. The main result. We can now define precisely what is meant by the informal statement that the operators d and ∂ are "dual" to each other. To understand the following statement, note that a k-form $\omega \in \Omega^k(M)$ induces a k-form $\Omega^k(L)$ on every submanifold $L \subset M$ by restriction, and this applies in particular to the boundary $\partial M \subset M$. Strictly speaking, the induced k-form on ∂M in this situation is $i^*\omega \in \Omega^k(\partial M)$ for the inclusion map $i : \partial M \hookrightarrow M$, but in the following we will also denote it by $\omega \in \Omega^k(\partial M)$ instead of $i^*\omega$.

THEOREM 12.11 (Stokes). *Assume M is an oriented n-manifold with boundary, where $n \geqslant 1$, and ∂M is equipped with its natural boundary orientation. Then for every $\omega \in \Omega_c^{n-1}(M)$,*

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

PROOF. As in the proof of Theorem 10.30, we can choose an open subset $M_0 \subset M$ with compact closure $\overline{M}_0$ such that $\text{supp}(\omega) \subset M_0$, and then choose a finite covering of $\overline{M}_0$ by oriented charts $\{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in I}$ and a partition of unity $\{\varphi_\alpha : M \to [0,1]\}$ such that each φ_α has compact support in $\mathcal{U}_\alpha$ and $\sum_{\alpha \in I} \varphi_\alpha \equiv 1$ on M_0. Then each $\omega_\alpha := \varphi_\alpha \omega$ belongs to $\Omega_c^{n-1}(\mathcal{U}_\alpha)$, and we have $\omega = \sum_{\alpha \in I} \omega_\alpha$ and $d\omega = \sum_{\alpha \in I} d\omega_\alpha$ on M_0. If we can then prove $\int_{\mathcal{U}_\alpha} d\omega_\alpha = \int_{\partial \mathcal{U}_\alpha} \omega_\alpha$ for each α, we will have

$$\int_M d\omega = \int_{M_0} d\omega = \sum_{\alpha \in I} \int_{M_0} d\omega_\alpha = \sum_{\alpha \in I} \int_{\mathcal{U}_\alpha} d\omega_\alpha = \sum_{\alpha \in I} \int_{\partial \mathcal{U}_\alpha} \omega_\alpha = \sum_{\alpha \in I} \int_{\partial M} \omega_\alpha = \int_{\partial M} \omega.$$

In this way, the problem has been reduced to the special case in which M is covered by a single chart.

Next, observe that if the theorem has been proven to hold on another oriented manifold N and there is an orientation-preserving diffeomorphism $\psi : M \to N$, then we can write $\omega = \psi^*\alpha$ for $\alpha := \psi_*\omega \in \Omega_c^{n-1}(N)$ and use Proposition 9.18 along with the invariance of the integral under pullbacks to conclude

$$\int_M d\omega = \int_M d(\psi^*\alpha) = \int_M \psi^*(d\alpha) = \int_N d\alpha = \int_{\partial N} \alpha = \int_{\partial M} \psi^*\alpha = \int_{\partial M} \omega,$$

where we have also used the fact that a diffeomorphism $M \to N$ necessarily maps ∂M to ∂N. The latter is true since diffeomorphisms between regions in $\mathbb{R}^n$ map open sets to open sets, and neighborhoods of boundary points in $\mathbb{H}^n$ are not open in $\mathbb{R}^n$.

The combined result of the previous two paragraphs is that it will suffice to prove Stokes' theorem in the case where M is an open subset $\mathcal{U} \subset \mathbb{H}^n$ in half-space; in fact, since we are going to assume $\omega \in \Omega_c^{n-1}(\mathcal{U})$ has compact support, we may as well also assume M is the whole half-space $\mathbb{H}^n$. The proof now becomes a simple computation based on Fubini's theorem and the fundamental theorem of calculus. We can write ω in terms of n compactly supported smooth functions $f_1, \ldots, f_n : \mathbb{H}^n \to \mathbb{R}$ as

$$\omega = f_i\, \alpha^i, \qquad \text{where} \qquad \alpha^i := dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n \in \Omega^{n-1}(\mathbb{H}^n),$$

and the hat indicates again that the corresponding term does *not* appear. Then $d\alpha^i = 0$ for each i, and $dx^j \wedge \alpha^i = 0$ for every $j \neq i$, thus

$$d\omega = df_i \wedge \alpha^i = \sum_{i=1}^n \partial_i f_i\, dx^i \wedge \alpha^i = \sum_{i=1}^n (-1)^{i-1} \partial_i f_i\, dx^1 \wedge \ldots \wedge dx^n,$$

where we have refrained from using the summation convention in the last two expressions in order to avert confusion. Of the n terms in this sum, we claim that $n-1$ of them vanish when integrated over $\mathbb{H}^n$. Let us check this specifically for $i = n$: choosing $N > 0$ large enough for the supports of the functions $f_1, \ldots, f_n$ to be contained in $[-N/2, 0] \times [-N/2, N/2]^{n-1}$, we use Fubini and the fundamental theorem of calculus to compute

$$\int_{\mathbb{H}^n} \partial_n f_n(x^1, \ldots, x^n)\, dx^1 \ldots dx^n = \int_{(-\infty,0] \times \mathbb{R}^{n-2}} \left(\int_{\mathbb{R}} \partial_n f_n(x^1, \ldots, x^n)\, dx^n \right) dx^1 \ldots dx^{n-1} = 0$$

since the assumption on the support of f_n implies

$$\begin{aligned} \int_{\mathbb{R}} \partial_n f_n(x^1, \ldots, x_n)\, dx^n &= \int_{-N}^{N} \partial_n f_n(x^1, \ldots, x_n)\, dx^n \\ &= f_n(x^1, \ldots, x^{n-1}, N) - f_n(x^1, \ldots, x^{n-1}, -N) = 0. \end{aligned}$$

This calculation works out the same way for each $i = 2, \ldots, n$, thus we find

$$\begin{aligned} \int_{\mathbb{H}^n} \omega &= \int_{\mathbb{H}^n} \partial_1 f_1(x^1, \ldots, x^n)\, dx^1 \ldots dx^n = \int_{\mathbb{R}^{n-1}} \left(\int_{(-\infty,0]} \partial_1 f_1(x^1, \ldots, x^n)\, dx^1 \right) dx^2 \ldots dx^n \\ &= \int_{\mathbb{R}^{n-1}} \left(\int_{-N}^{0} \partial_1 f_1(x^1, \ldots, x^n)\, dx^1 \right) dx^2 \ldots dx^n \\ &= \int_{\mathbb{R}^{n-1}} \left(f_1(0, x^2, \ldots, x^n) - f_1(-N, x^2, \ldots, x^n) \right) dx^2 \ldots dx^n \\ &= \int_{\mathbb{R}^{n-1}} f_1(0, x^2, \ldots, x^n)\, dx^2 \ldots dx^n = \int_{\partial \mathbb{H}^n} f_1 \; dx^2 \wedge \ldots \wedge dx^n. \end{aligned}$$

This last expression is $\int_{\partial \mathbb{H}^n} \omega$, as all other terms in ω contain dx^1, which vanishes when restricted to $\partial \mathbb{H}^n$. □

EXAMPLE 12.12. For a smooth function $f : [a, b] \to \mathbb{R}$ on a nontrivial compact interval, we can denote the standard coordinate on $\mathbb{R}$ by x and write $df = f'\, dx$. The fundamental theorem of calculus then amounts to the following special case of Stokes' theorem,

$$\int_a^b f'(x)\, dx = \int_{[a,b]} df = \int_{-\{a\} \amalg \{b\}} f = f(b) - f(a).$$

With this example in mind, Stokes' theorem is considered to be the natural n-dimensional generalization of the fundamental theorem of calculus.

EXERCISE 12.13. Prove the following version of *integration by parts*: if M is a compact oriented n-manifold with boundary, $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^\ell(M)$ with $k + \ell = n - 1$, then

$$\int_M d\alpha \wedge \beta = \int_{\partial M} \alpha \wedge \beta - (-1)^k \int_M \alpha \wedge d\beta.$$

EXAMPLE 12.14. Heuristically, the discussion of §12.3 suggests that if M and N are compact manifolds with boundary having dimensions m and n respectively, then for any $\omega \in \Omega^{m+n-1}(M \times N)$, one should have

$$\int_{M \times N} d\omega = \int_{\partial M \times N} \omega + (-1)^m \int_{M \times \partial N} \omega. \tag{12.4}$$

Here the right hand side is obtained from the integral of ω over $\partial(M \times N)$ by splitting the latter into the two *almost* disjoint subsets $\partial M \times N$ and $M \times \partial N$ (whose intersection $\partial M \times \partial N$ is a set of measure zero in either one), and then including a sign (cf. Exercise 11.4) to account for the fact that the product orientation of $M \times \partial N$ only matches the boundary orientation of $\partial(M \times N)$ when m is odd. As it stands, the left hand side of (12.4) does not immediately make sense unless either ∂M or ∂N is empty (in which case (12.4) follows from Stokes' theorem), because $M \times N$ is otherwise not a smooth manifold with boundary. There are at least two ways that one could nonetheless make sense of (12.4):

(1) Define the notion of an oriented manifold *with boundary and corners* by allowing open subsets of $(-\infty, 0]^2 \times \mathbb{R}^{n-2}$ as local coordinate models, generalize the definition of the integral to this wider class of manifolds and prove that Stokes' theorem still holds if $\partial(M \times N)$ is understood in the sense of §12.3. This requires a bit of extra bookkeeping, but is not fundamentally more difficult than what we have already done.
(2) Choose a nested sequence of closed subsets $A_1 \subset A_2 \subset \ldots \bigcup_{j \in \mathbb{N}} A_j = M \times N$ such that each A_j is a smooth manifold with boundary (obtained by "smoothing the corner" of $M \times N$ in progressively small neighborhoods of $\partial M \times \partial N$), then define $\int_{M \times N} d\omega$ to mean $\lim_{j \to \infty} \int_{A_j} d\omega$ and deduce (12.4) from $\int_{A_j} d\omega = \int_{\partial A_j} \omega$.

REMARK 12.15. Much time and effort has been wasted by well-intentioned mathematicians trying to determine whether the correct orthography should be "Stokes' theorem" or "Stokes's theorem". After a years-long struggle I came to the conclusion that it is, essentially, a matter of personal taste. What I can say with absolute certainty is that it is not "Stoke's theorem".

12.5. The classical integration theorems. Various results that are considered central in classical vector calculus are easy consequences of Stokes' theorem.

12.5.1. *Divergence.* The **divergence** (*Divergenz*) of a vector field $X \in \mathfrak{X}(M)$ with respect to a volume form $d\text{vol} \in \Omega^n(M)$ is defined as the unique real-valued function $\text{div}(X) : M \to \mathbb{R}$ such that

$$d(\iota_X d\text{vol}) = \text{div}(X) \cdot d\text{vol}. \tag{12.5}$$

The definition makes sense because $\iota_X d\text{vol}$ is an $(n-1)$-form and thus $d(\iota_X d\text{vol})$ is an n-form, and every n-form is at each point a scalar multiple of the given volume form. It may not seem obvious at this stage why $\text{div}(X)$ is a natural thing to define—we will address this question more thoroughly next week—but the following exercise should at least make it look familiar.

EXERCISE 12.16. Assume M is an n-manifold with a fixed volume form $d\mathrm{vol} \in \Omega^n(M)$, $(\mathcal{U}, x)$ is a chart on M and $f : \mathcal{U} \to \mathbb{R}$ is the unique function such that $d\mathrm{vol} = f\,dx^1 \wedge \ldots \wedge dx^n$ on $\mathcal{U}$. Show that for any $X \in \mathfrak{X}(M)$,

$$\mathrm{div}(X) = \frac{1}{f}\partial_i(fX^i) \qquad \text{on } \mathcal{U}.$$

In particular for the standard volume form $d\mathrm{vol} = dx^1 \wedge \ldots \wedge dx^n$ on $\mathbb{R}^n$, this reduces to the standard definition of divergence in vector calculus.

If M is a compact oriented n-manifold with boundary carrying a positive volume form $d\mathrm{vol}_M \in \Omega^n(M)$ and $X \in \mathfrak{X}(M)$ is a vector field, Stokes' theorem now implies

$$\int_M \mathrm{div}(X)\,d\mathrm{vol}_M = \int_M d(\iota_X d\mathrm{vol}_M) = \int_{\partial M} \iota_X d\mathrm{vol}_M. \tag{12.6}$$

The geometric meaning of this last integral is best understood in the special case where $d\mathrm{vol}_M$ is the Riemannian volume form compatible with a Riemannian metric g on M, which we shall write in the following using the usual notation for inner products,

$$\langle X, Y\rangle := g(X,Y) \qquad \text{for } X, Y \in T_pM,\ p \in M.$$

By Proposition 11.14, the Riemannian volume form $d\mathrm{vol}_{\partial M}$ on ∂M is then

$$d\mathrm{vol}_{\partial M} := \iota_\nu d\mathrm{vol}_M|_{T(\partial M)} \in \Omega^{n-1}(\partial M),$$

where ν is the unique outward-pointing normal vector field to ∂M. (You should take a moment to convince yourself that we are getting the orientations right, i.e. $d\mathrm{vol}_{\partial M}$ really is a *positive* volume form with respect to the boundary orientation of ∂M.) To relate this to $\iota_X d\mathrm{vol}_M$, observe that along ∂M, $X = \langle X, \nu\rangle \nu + Y$ for a unique vector field $Y \in \mathfrak{X}(\partial M)$, but $\iota_Y d\mathrm{vol}_M$ vanishes when restricted to the boundary because feeding it any $(n-1)$-tuple of vectors $Y_1, \ldots, Y_{n-1}$ tangent to ∂M means evaluating $d\mathrm{vol}_M$ on $(Y, Y_1, \ldots, Y_{n-1})$, and those are *all* tangent to the $(n-1)$-dimensional boundary and thus cannot be linearly independent. We conclude

$$\iota_X d\mathrm{vol}_M|_{T(\partial M)} = \langle X, \nu\rangle\ \iota_\nu d\mathrm{vol}_M|_{T(\partial M)} = \langle X, \nu\rangle\, d\mathrm{vol}_{\partial M},$$

and the implication of (12.6) is thus

$$\boxed{\int_M \mathrm{div}(X)\,d\mathrm{vol}_M = \int_{\partial M} \langle X, \nu\rangle\, d\mathrm{vol}_{\partial M}}. \tag{12.7}$$

This is a mild generalization of the classical result known as *Gauss's divergence theorem.*[43] Physics textbooks like to write their favorite special case of this result in some form such as

$$\iiint_\Omega (\nabla \cdot \mathbf{X})\,dV = \oiint_{\partial\Omega} \mathbf{X} \cdot d\mathbf{a}, \tag{12.8}$$

where $\Omega \subset \mathbb{R}^3$ is assumed to be a compact region bounded by a smooth surface $\partial\Omega \subset \mathbb{R}^3$, $\nabla \cdot \mathbf{X}$ is the divergence of a vector field $\mathbf{X} \in \mathfrak{X}(\Omega)$ with respect to the standard volume form $d\mathrm{vol}_{\mathbb{R}^3} := dx \wedge dy \wedge dz$, the "V" in $dV := d\mathrm{vol}_{\mathbb{R}^3}$ stands for "volume" and the "a" in $\mathbf{X} \cdot d\mathbf{a} := \langle \mathbf{X}, \nu\rangle\, d\mathrm{vol}_{\partial\Omega}$ stands for "area". (The symbol $d\mathbf{a}$ in this situation is thought of as a "vector-valued measure" that encodes not only the 2-dimensional measure on $\partial\Omega$ but also its normal vector field.) The repetition of the integral signs corresponds to the dimension of the manifold and can be seen as a reference to Fubini's theorem; the additional loop in $\oiint$ merely refers to the fact that $\partial\Omega$ is a "closed" surface (the 2-dimensional analogue of a closed loop), i.e. it is compact and has no boundary. Gauss's theorem has an important interpretation in electrostatics: if $\mathbf{X}$ represents the electric field on a

[43] or possibly "Gauss' divergence theorem", I don't know

region $\Omega \subset \mathbb{R}^3$, then its divergence is the electrical charge density, and (12.8) thus says that the total electrical charge in the region Ω is equal to the total *flux* of the electric field through the boundary of Ω.

12.5.2. *Curl.* The next example only makes sense in the case

$$\dim M = 3.$$

It relies on the observation that for any n-dimensional vector space V with a nontrivial top-dimensional form $\omega \in \Lambda^n V^*$, the map

$$V \to \Lambda^{n-1} V^* : v \mapsto \iota_v \omega$$

is an isomorphism. Indeed, it is clearly injective since $\omega \neq 0$ and any $v \neq 0$ can be extended to a basis of V, so surjectivity then follows from the fact that $\dim \Lambda^{n-1} V^* = \binom{n}{n-1} = n = \dim V$. With this understood, any volume form $d\mathrm{vol}_M$ on a 3-manifold M determines an isomorphism

$$\mathfrak{X}(M) \xrightarrow{\cong} \Omega^2(M) : X \mapsto \iota_X d\mathrm{vol}_M.$$

Let us now assume (M, g) is an oriented Riemannian 3-manifold and $d\mathrm{vol}_M$ is its Riemannian volume form. The metric $\langle\ ,\ \rangle := g$ also determines an isomorphism

$$\mathfrak{X}(M) \xrightarrow{\cong} \Omega^1(M) : X \mapsto X_\flat := \langle X, \cdot \rangle.$$

The **curl** (*Rotation*) of $X \in \mathfrak{X}(M)$ is then defined as the unique vector field $\mathrm{curl}(X) \in \mathfrak{X}(M)$ such that

$$\iota_{\mathrm{curl}(X)} d\mathrm{vol}_M = d(X_\flat).$$

EXERCISE 12.17. Convince yourself that on $M := \mathbb{R}^3$ with its standard Riemannian metric defined via the Euclidean inner product, the curl of a vector field is the same thing that you learned about once upon a time in vector calculus.

Now if $\Sigma \subset M$ is an oriented 2-dimensional submanifold with boundary, Σ and $\partial\Sigma$ each inherit Riemannian metrics as submanifolds of M, and thus have canonical Riemannian volume forms $d\mathrm{vol}_\Sigma$ and $d\mathrm{vol}_{\partial\Sigma}$ respectively. For an appropriate choice[44] of normal vector field ν along Σ, Proposition 11.14 implies

$$d\mathrm{vol}_\Sigma = \iota_\nu d\mathrm{vol}_M|_{T\Sigma} \in \Omega^2(\Sigma),$$

and a repeat of the same argument we used for the divergence theorem then implies that for any $Y \in \mathfrak{X}(M)$,

$$\iota_Y d\mathrm{vol}_M|_{T\Sigma} = \langle Y, \nu \rangle\, d\mathrm{vol}_\Sigma.$$

If $Y = \mathrm{curl}(X)$ for some $X \in \mathfrak{X}(M)$, Stokes' theorem now implies

$$\int_\Sigma \langle \mathrm{curl}(X), \nu \rangle\, d\mathrm{vol}_\Sigma = \int_\Sigma d(X_\flat) = \int_{\partial\Sigma} X_\flat.$$

To understand the integral on the right, let $\tau \in \mathfrak{X}(\partial\Sigma)$ denote the unique positively-oriented unit vector field on $\partial\Sigma$, so $d\mathrm{vol}_{\partial\Sigma}(\tau) = 1$, and $X_\flat(\tau) = \langle X, \tau \rangle$ thus implies $X_\flat|_{T(\partial\Sigma)} = \langle X, \tau \rangle\, d\mathrm{vol}_{\partial\Sigma}$, and we obtain

$$\boxed{\int_\Sigma \langle \mathrm{curl}(X), \nu \rangle\, d\mathrm{vol}_\Sigma = \int_{\partial\Sigma} \langle X, \tau \rangle\, d\mathrm{vol}_{\partial\Sigma}}. \tag{12.9}$$

[44]One can deduce from the assumption that both M and Σ are oriented that a normal vector field ν along Σ exists, and there are multiple choices—if Σ is connected, then there are exactly two choices, differing by a sign. The *appropriate* choice is the one that makes the volume form $\iota_\nu d\mathrm{vol}_M$ on Σ positive.

This generalizes what is usually called the "classical" Stokes' theorem in vector calculus. In physics textbooks, one finds it written for the case $\Sigma \subset \mathbb{R}^3$ with the standard metric as

$$\iint_\Sigma (\nabla \times \mathbf{X}) \cdot d\mathbf{a} = \oint_{\partial\Sigma} \mathbf{X} \cdot d\mathbf{l},$$

where $\nabla \times \mathbf{X}$ denotes the curl of $\mathbf{X} \in \mathfrak{X}(\mathbb{R}^3)$, $d\mathbf{a}$ is the same "vector-valued measure" that appeared in (12.8), and $d\mathbf{l}$ similarly denotes a 1-dimensional vector-valued measure that encodes both the volume form $d\mathrm{vol}_{\partial\Sigma}$ and the tangent vector field τ.

13. Closed and exact forms

13.1. Some easy applications of Stokes. The following terminology is used consistently throughout differential geometry.

Definition 13.1. A manifold M is **closed** (*geschlossen*) if it is compact and $\partial M = \varnothing$. We say that M is **open** (*offen*) if none of its connected components are closed, i.e. they all are noncompact and/or have nonempty boundary.[45]

Example 13.2. Manifolds of dimension 0 never have boundary, so a 0-manifold is closed if and only if it is compact, i.e. it is a discrete finite set.

Example 13.3. If M is a compact manifold with boundary, then ∂M is a closed manifold.

Definition 13.4. A differential form $\omega \in \Omega^k(M)$ is called **closed** (*geschlossen*) if $d\omega = 0$, and it is called **exact** (*exakt*) of $\omega = d\alpha$ for some $\alpha \in \Omega^{k-1}(M)$. In the latter situation, the form α is called a **primitive** of ω.

Example 13.5. A closed 0-form is the same thing as a locally constant function, and an exact 1-form is the same thing as a differential. There are no exact 0-forms since there is no such thing as a (-1)-form.

Example 13.6. On an n-manifold, every n-form is closed since there are no nontrivial $(n+1)$-forms.

Example 13.7. Given a volume form $d\mathrm{vol} \in \Omega^n(M)$, a vector field $X \in \mathfrak{X}(M)$ has vanishing divergence if and only if the $(n-1)$-form $\iota_X d\mathrm{vol}$ is closed. Similarly, if (M, g) is an oriented Riemannian 3-manifold, $X \in \mathfrak{X}(M)$ has vanishing curl if and only if the 1-form $X_\flat := g(X, \cdot)$ is closed.

Here is a bit of low-hanging fruit that can be picked as soon as one understands the above definitions and the statement of Stokes' theorem.

Proposition 13.8. *If M is a closed oriented n-manifold and $\omega \in \Omega^n(M)$ is exact, then $\int_M \omega = 0$. Similarly, if M is a compact oriented n-manifold with boundary and $\alpha \in \Omega^{n-1}(M)$ is closed, then $\int_{\partial M} \alpha = 0$.*

---

[45] Be aware that the word "closed" has a different meaning when referring to a manifold than it does when referring to a subset of a topological space. For instance, if M is a manifold, then a compact submanifold $\Sigma \subset M$ with boundary is a closed subset of M, but it is not a closed manifold if $\partial\Sigma \neq \varnothing$. The German language uses two different words for these separate meanings of "closed": a subset in a topological space can be *abgeschlossen*, but a manifold can be *geschlossen*.

PROOF. If you review the proof of Stokes' theorem, you will find that it is valid in the case $\partial M = \varnothing$ so long as one understands every integral over $\varnothing$ to be 0 by definition. Thus $\partial M = \varnothing$ and $\omega = d\beta$ for some $\beta \in \Omega^{n-1}(M)$ implies

$$\int_M \omega = \int_M d\beta = \int_\varnothing \beta = 0,$$

and if ∂M is not assumed empty but $\alpha \in \Omega^{n-1}(M)$ is closed,

$$\int_{\partial M} \alpha = \int_M d\alpha = 0.$$

□

COROLLARY 13.9. *On a closed oriented n-manifold M, every n-form $\omega \in \Omega^n(M)$ with $\int_M \omega \neq 0$ is closed but not exact. In particular, this is true whenever ω is a volume form.* □

REMARK 13.10. One can show that Corollary 13.9 fails whenever either $\partial M \neq \varnothing$ or M is noncompact. In the former case, $\int_M \omega \neq 0$ for an exact form $\omega = d\alpha$ is not a contradiction, since $\int_{\partial M} \alpha$ might also be nonzero. There is a different problem if M has empty boundary but is noncompact: the use of Stokes' theorem to derive the contradiction $0 \neq \int_M d\alpha = \int_{\partial M} \alpha = 0$ is not valid unless α has compact support, so it can happen for instance that $\omega \in \Omega^n_c(M)$ satisfies $\int_M \omega \neq 0$ and is the exterior derivative of an $(n-1)$-form whose support is noncompact. We will see shortly that, indeed, *every* n-form on $\mathbb{R}^n$ for $n \geqslant 1$ is exact (see Corollary 13.34 below).

EXERCISE 13.11. Show that for each $k \geqslant 0$, a k-form $\omega \in \Omega^k(M)$ is closed if and only for every compact oriented $(k+1)$-dimensional submanifold $L \subset M$ with boundary, $\int_{\partial L} \omega = 0$.
Hint: For any point $p \in M$ and linearly-independent vectors $X_1, \ldots, X_{k+1} \in T_pM$, you could choose $L \subset M$ to be a small $(k+1)$-disk through p tangent to the space spanned by $X_1, \ldots, X_{k+1}$.

13.2. The Poincaré lemma and simple connectedness. The observation in Example 13.3 that boundaries of compact manifolds are closed has a dual statement for differential forms: since $d^2 := d \circ d = 0$, every exact differential form is also closed. Corollary 13.9 reveals however that the converse is generally false. Here is a more concrete example.

EXAMPLE 13.12. On $\mathbb{R}^2 \backslash \{0\}$, one can define a smooth 1-form in Cartesian coordinates (x, y) by

$$\lambda := \frac{1}{x^2 + y^2} (x\,dy - y\,dx).$$

This expression takes a more revealing form of one rewrites it in polar coordinates: assume $\mathcal{U} \subset \mathbb{R}^2 \backslash \{0\}$ is a subset on which there is a well-defined chart of the form $(r, \theta) : \mathcal{U} \to \mathbb{R}^2$ such that r takes positive values and the relations $x = r\cos\theta$ and $y = r\sin\theta$ hold; concretely, we can take $\mathcal{U}$ to be the complement of a ray $\{tv \in \mathbb{R}^2 \mid t \in [0, \infty)\}$ for some $v \in \mathbb{R}^2 \backslash \{0\}$, and the image of θ is then an open interval of the form $(c, c + 2\pi)$. In terms of r and θ, we have $dx = (\cos\theta)\,dr - (r\sin\theta)\,d\theta$ and $dy = (\sin\theta)\,dr + (r\cos\theta)\,d\theta$, thus

$$\lambda = \frac{1}{r^2} \left[r\cos\theta\,(\sin\theta\,dr + r\cos\theta\,d\theta) - r\sin\theta\,(\cos\theta\,dr - r\sin\theta\,d\theta) \right] = d\theta,$$

so λ is exact on $\mathcal{U}$. Since this computation holds independently of the choice of domain $\mathcal{U} \subset \mathbb{R}^2 \backslash \{0\}$, it follows that $d\lambda = 0$ everywhere. But the restriction of $(\mathcal{U}, (r, \theta))$ to $\{r = 1\}$ now defines a chart on $S^1 \subset \mathbb{R}^2 \backslash \{0\}$ in the form $(S^1 \backslash \{q\}, \theta)$ for some point $q \in S^1$, which is a set of measure zero, thus $\int_{S^1} \lambda$ can be computed using the methods of §11.2, and the answer is

$$\int_{S^1} \lambda = \int_{(c, c+2\pi)} d\theta = 2\pi \neq 0.$$

This clearly could not happen if λ were df for some $f \in \Omega^0(\mathbb{R}^2\backslash\{0\}) = C^\infty(\mathbb{R}^2\backslash\{0\})$, as the restriction of λ to S^1 would then be $d(f|_{S^1})$ and we would have a contradiction to Proposition 13.8.

REMARK 13.13. It is conventional to denote the 1-form in Example 13.12 by

$$d\theta \in \Omega^1(\mathbb{R}^2\backslash\{0\})$$

even though, strictly speaking, it is not the differential of any smooth function $\theta : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}$. One reasonable way to think about it is that while θ cannot be defined on this domain as a smooth *real*-valued function, it can be defined to take values in the quotient $\mathbb{R}/2\pi\mathbb{Z}$, which is a smooth manifold and $\theta : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}/2\pi\mathbb{Z}$ in this sense is a smooth map. The latter means in practice that any point $p \in \mathbb{R}^2\backslash\{0\}$ admits a neighborhood $\mathcal{U} \subset \mathbb{R}^2\backslash\{0\}$ on which the smooth function $\theta : \mathcal{U} \to \mathbb{R}$ can be defined, though this function is not unique, as it can equally well be replaced by $\theta + 2\pi m$ for any $m \in \mathbb{Z}$. But modifying θ by addition of a constant does not change its differential, thus $d\theta$ is uniquely defined.

Remark 13.13 illustrates a phenomenon that is generalized in the following result: every closed differential form is "locally" exact.

THEOREM 13.14 (the Poincaré Lemma). *If $\omega \in \Omega^k(M)$ is closed and $k \geqslant 1$, then for every $p \in M$ there exists a neighborhood $\mathcal{U} \subset M$ of p and a $(k-1)$-form $\alpha \in \Omega^{k-1}(\mathcal{U})$ such that $d\alpha = \omega$ on $\mathcal{U}$.*

A proof of the Poincaré lemma will be given at the end of this lecture. The next two results are easier to prove, but imply a stronger statement for the case $k = 1$.

LEMMA 13.15. *A 1-form $\lambda \in \Omega^1(M)$ is exact if and only if $\int_{S^1} \gamma^*\lambda = 0$ for all smooth maps $\gamma : S^1 \to M$.*

PROOF. If $\lambda = df$ for some $f \in C^\infty(M)$, then Proposition 13.8 implies $\int_{S^1} \gamma^*\lambda = \int_{S^1} \gamma^* df = \int_{S^1} d(\gamma^* f) = 0$ for every smooth map $\gamma : S^1 \to M$. Conversely, assume $\int_{S^1} \gamma^*\lambda$ always vanishes. The following recipe for constructing a function $f : M \to \mathbb{R}$ with $df = \lambda$ can be applied on every connected component of M separately, so we may as well assume M is connected. We claim that if we fix a reference point $p_0 \in M$, then $f : M \to \mathbb{R}$ can be defined by

$$f(p) := \int_0^a \lambda(\dot\gamma(t))\, dt \quad \text{for any } a > 0,\ \gamma \in C^\infty([0,a], M) \text{ with } \gamma(0) = p_0,\ \gamma(a) = p. \tag{13.1}$$

We must first show that $f(p)$ is independent of the choice of the path $\gamma : [0,a] \to M$ from p_0 to p. To this end, here are two useful observations: first, by the substitution rule, the integral in (13.1) does not change if we replace $\gamma : [0,a] \to M$ with $\gamma \circ \psi : [0,1] \to M$ for any smooth map $\psi : [0,1] \to [0,a]$ with $\psi(0) = 0$ and $\psi(1) = a$. As a consequence, we lose no generality by restricting our attention to paths $\gamma : [0,1] \to M$ that are constant on neighborhoods of 0 and 1, with values p_0 and p respectively. The second observation is that if t denotes the standard coordinate on the 1-manifold $[0,1] \subset \mathbb{R}$, then $(\gamma^*\lambda)_t(\partial_t) = \lambda_{\gamma(t)}(\gamma_*\partial_t) = \lambda_{\gamma(t)}(\dot\gamma(t))$, thus we can also write

$$f(p) = \int_{[0,1]} \gamma^*\lambda.$$

Now if $\gamma_1, \gamma_2 : [0,1] \to M$ are two smooth paths from p_0 to p that are both constant near 0 and 1, we can concatenate γ_1 with the reversal of γ_2 to form a smooth loop $\varphi : S^1 \to M$ in the form

$$\varphi(e^{\pi i t}) = \begin{cases} \gamma_1(t) & \text{for } 0 \leqslant t \leqslant 1, \\ \gamma_2(2-t) & \text{for } 1 \leqslant t \leqslant 2, \end{cases}$$

where for convenience we are identifying $\mathbb{R}^2$ in the obvious way with $\mathbb{C}$ so that $S^1 \subset \mathbb{C}$. If we now split S^1 into its upper and lower semicircles $S^1_\pm$ with parametrizations $\psi_\pm : [0,1] \to S^1_\pm : t \mapsto e^{\pi i t}$, we have $\gamma_1 = \varphi \circ \psi_+$ and $\gamma_2 = \varphi \circ \psi_-$, but ψ_+ is orientation preserving while ψ_- is orientation reversing, thus

$$0 = \int_{S^1} \varphi^*\lambda = \int_{S^1_+} \varphi^*\lambda + \int_{S^1_-} \varphi^*\lambda = \int_{\psi_+([0,1])} \varphi^*\lambda + \int_{\psi_-([0,1])} \varphi^*\lambda$$
$$= \int_{[0,1]} \psi_+^*\varphi^*\lambda - \int_{[0,1]} \psi_-^*\varphi^*\lambda = \int_{[0,1]} (\varphi \circ \psi_+)^*\lambda - \int_{[0,1]} (\varphi \circ \psi_-)^*\lambda = \int_{[0,1]} \gamma_1^*\lambda - \int_{[0,1]} \gamma_2^*\lambda.$$

With independence of the choice of γ established, we observe that (13.1) implies $\frac{d}{dt} f(\gamma(t)) = \lambda(\dot\gamma(t))$ for every t and every smooth path γ starting at p_0, thus $df = \lambda$. □

Exercise 13.16. Use a slight modification of the proof of Lemma 13.15 to show that on S^1, a 1-form $\lambda \in \Omega^1(S^1)$ is exact if and only if $\int_{S^1} \lambda = 0$.

Definition 13.17. A smooth manifold M is **simply connected** (*einfach zusammenhängend*) if it is connected and every smooth map $\gamma : S^1 \to M$ admits a smooth extension over the 2-disk, i.e. a map $u : \mathbb{D}^2 \to M$ such that $u|_{\partial \mathbb{D}^2} = \gamma$.

Remark 13.18. In algebraic topology, a topological space is called simply connected if it is path-connected and its fundamental group vanishes, but for smooth manifolds, Definition 13.17 is equivalent to this condition. In particular, one could replace the word "smooth" by "continuous" without changing anything, because by general perturbation results in differential topology (see e.g. [**Hir94**]), continuous maps between smooth manifolds always admit smooth approximations.

Theorem 13.19. *If M is a simply connected manifold, then every closed 1-form $\lambda \in \Omega^1(M)$ is exact.*

Proof. If $\lambda \in \Omega^1(M)$ is closed and every smooth map $\gamma : S^1 \to M$ admits a smooth extension $u : \mathbb{D}^2 \to M$, then

$$\int_{S^1} \gamma^*\lambda = \int_{\partial \mathbb{D}^2} u^*\lambda = \int_{\mathbb{D}^2} d(u^*\lambda) = \int_{\mathbb{D}^2} u^*(d\lambda) = 0,$$

hence λ satisfies the criterion of Lemma 13.15 and is therefore exact. □

It should be easy to convince yourself that every convex subset of $\mathbb{R}^n$ is simply connected, and every point in a manifold has a neighborhood that looks like a convex subset of $\mathbb{R}^n$ in local coordinates, implying in turn that that neighborhood is simply connected. Theorem 13.19 thus implies the $k = 1$ case of the Poincaré lemma. But it also implies more, because there are many simply connected manifolds that are more interesting than convex sets.

Example 13.20. For each $n \geqslant 2$, the sphere S^n is simply connected. Here is an incomplete but (maybe?) believable proof: since $\dim S^n > \dim S^1$, no smooth map $\gamma : S^1 \to S^n$ can be surjective,[46] i.e. it must miss at least one point $p \in S^n$ and can thus be viewed as a map $S^1 \to S^n \backslash \{p\}$. But by stereographic projection, one can also find a diffeomorphism of $S^n \backslash \{p\}$ to $\mathbb{R}^n$ and then appeal to the fact that $\mathbb{R}^n$ (as a convex set) is simply connected. It follows that closed 1-forms on S^n for $n \geqslant 2$ are always exact.

[46] I'm pretty sure that you cannot visualize any surjective smooth map $f : M \to N$ when $\dim M < \dim N$, though actually proving they don't exist is not completely trivial. It follows easily from Sard's theorem, a fundamental result in differential topology stating that the set of critical values of a smooth map $f : M \to N$ always has measure zero. This means that for almost every $q \in N$, $T_p f : T_p M \to T_q N$ is surjective for every $p \in f^{-1}(q)$; the only way for this to hold when $\dim M < \dim N$ is if $f^{-1}(q) = \emptyset$. The much more surprising fact is that *continuous* maps $f : M \to N$ can be surjective, even when $\dim N > \dim M$; look up the term "space-filling curve". Such maps can never be smooth.

Remark 13.21. You may have noticed that in Theorem 13.19, it would have sufficed to assume that every smooth map $\gamma : S^1 \to M$ admits a smooth extension $u : \Sigma \to M$ over *some* compact, smooth, oriented surface Σ with boundary $\partial\Sigma = S^1$, i.e. not necessarily the disk, but any surface whose boundary is a circle. (An easy example would be obtained by cutting a hole out of the 2-torus $\mathbb{T}^2$.) This means that Theorem 13.19 is true under a somewhat more general hypothesis than simple connectedness. The natural language for this generalization is homology, i.e. the theorem holds for any manifold M whose first homology group with real coefficients vanishes. A full explanation of this statement would require a major digression into algebraic topology, so we will not discuss it any further here, but suffice it to say that in dimension 2, there are no examples for which this distinction makes a difference, but in dimension 3 there are. Poincaré famously conjectured that every closed 3-manifold with vanishing first homology group is homeomorphic to S^3, but later found an example—now known as the *Poincaré homology sphere*—that satisfies this hypothesis but (unlike S^3) is not simply connected, and thus had to revise his conjecture. The revised conjecture was proved over 100 years later.

Example 13.22. On a Riemannian manifold (M, g), the inner product $\langle\ ,\ \rangle := g$ determines an isomorphism $T_pM \to T_p^*M : X \mapsto X_\flat := \langle X, \cdot\rangle$ at every point $p \in M$, which can be used to associate to any smooth function $f : M \to \mathbb{R}$ its **gradient** vector field $\nabla f \in \mathfrak{X}(M)$, uniquely determined by

$$df = \langle\nabla f, \cdot\rangle.$$

A vector field $X \in \mathfrak{X}(M)$ cannot be the gradient of a function unless the 1-form $X_\flat \in \Omega^1(M)$ is closed, and conversely, the Poincaré lemma implies that every vector field satisfying this condition is *locally* the gradient of a function, though perhaps not globally (unless M is simply connected). If M is oriented and 3-dimensional, then this result can also be expressed in terms of the curl (cf. §12.5.2): any gradient $X = \nabla f$ satisfies $\iota_{\mathrm{curl}(X)} d\mathrm{vol}_M = d(df) = 0$, implying

$$\mathrm{curl}(\nabla f) \equiv 0,$$

and conversely, any vector field $X \in \mathfrak{X}(M)$ with $\mathrm{curl}(X) \equiv 0$ is locally the gradient of a function.

In the same context, the curl of any vector field $X \in \mathfrak{X}(M)$ satisfies $\iota_{\mathrm{curl}(X)} d\mathrm{vol}_M = d(X_\flat)$ and thus $d(\iota_{\mathrm{curl}(X)} d\mathrm{vol}_M) = d^2(X_\flat) = 0$, implying

$$\mathrm{div}(\mathrm{curl}(X)) \equiv 0.$$

Conversely, any divergenceless vector field $Y \in \mathfrak{X}(M)$ satisfies $d(\iota_Y d\mathrm{vol}_M) = 0$, so that by the Poincaré lemma, $\iota_Y d\mathrm{vol}_M \in \Omega^2(M)$ can be written on any sufficienly small neighborhood $\mathcal{U}$ as $d\lambda$ for some $\lambda \in \Omega^1(\mathcal{U})$. The latter is also $X_\flat$ for a unique vector field $X \in \mathfrak{X}(\mathcal{U})$, whose curl is therefore Y: in other words, any divergenceless vector field is locally the curl of another vector field.

While (13.1) provides a fairly straightforward recipe to find a local primitive of any closed 1-form, it is not as easy to derive local primitives for closed k-forms when $k \geqslant 2$. One possible approach is to work on "boxes" of the form $M := (a_1, b_1) \times \ldots \times (a_n, b_n)$ and proceed by induction on the number of dimensions, showing that if one can already find primitives for closed k-forms on the hypersurface $\Sigma_c := (a_1, b_1) \times \ldots \times (a_{n-1}, b_{n-1}) \times \{c\}$ for some constant $c \in (a_n, b_n)$, then primitives on Σ_c can be extended to primitives on M by integrating in the nth direction. I have proved the Poincaré lemma in this way when I've taught analysis courses (see [Wen19]), but the idea behind the argument has a tendency to get lost behind computational details. We will adopt a different approach in these notes, and deduce the Poincaré lemma from a deeper theorem about the homotopy-invariance of de Rham cohomology. We will see at the end that this approach does lead to an explicit formula generalizing (13.1) to produce local primitives of closed k-forms (see in partiular Remark 13.39), but in contrast with (13.1), one would be very unlikely to find this formula from an educated guess.

13.3. De Rham cohomology. By now we have gathered some evidence that the distinction between closed and exact forms on a manifold M has something to do with the topology of M. We shall now formalize this relation by defining an algebraic invariant of smooth manifolds.

Definition 13.23. For a smooth n-manifold M and each integer $k \in \mathbb{Z}$, let $d_k : \Omega^k(M) \to \Omega^{k+1}(M)$ denote the restriction of the exterior derivative $d : \Omega^*(M) \to \Omega^*(M)$ to the subspace $\Omega^k(M) \subset \Omega^*(M)$, with the convention that for $k < 0$, $\Omega^k(M)$ is the trivial subspace (hence d_{-1} is the trivial map into $\Omega^0(M)$). The k**th de Rham cohomology** of M is the vector space

$$H^k_{\mathrm{dR}}(M) := \ker(d_k)\big/ \operatorname{im}(d_{k-1}),$$

i.e. it is the quotient of the space of closed k-forms by the subspace of exact k-forms. We write

$$H^*_{\mathrm{dR}}(M) := \bigoplus_{k\in\mathbb{Z}} H^k_{\mathrm{dR}}(M).$$

Remark 13.24. The case $k < 0$ was included in Definition 13.23 only in order to make sure that the definition of $H^0_{\mathrm{dR}}(M)$ makes sense, but $H^k_{\mathrm{dR}}(M)$ for $k < 0$ is just the trivial vector space, and we will have no need to mention it again. It is similarly easy to see that $H^k_{\mathrm{dR}}(M) = 0$ whenever $k > \dim M$, since the space of k-forms is already trivial in this case. Thus in practice, $H^k_{\mathrm{dR}}(M)$ is potentially interesting only for k in the range $0 \leqslant k \leqslant \dim M$.

It may seem surprising at first glance that $H^k_{\mathrm{dR}}(M)$ is useful or computable: in typical cases both $\ker(d_k)$ and $\operatorname{im}(d_{k-1})$ are infinite-dimensional vector spaces, and one would not normally expect the quotient of one infinite-dimensional space by another one to carry interesting information. It turns out however that in almost all interesting cases, the quotient is finite dimensional, and its dimension is a useful numerical invariant of manifolds. Let us first clarify what is meant by the word "invariant".

Proposition 13.25. *For smooth maps $f : M \to N$, the linear map $f^* : \Omega^k(N) \to \Omega^k(M)$ sends closed forms on N to closed forms on M, and it also descends*[47] *to the quotients to define a linear map $f^* : H^k_{\mathrm{dR}}(N) \to H^k_{\mathrm{dR}}(M)$ that satisfies the following properties:*

(1) For another smooth map $g : N \to Q$, $(g \circ f)^* = f^* g^* : H^k_{\mathrm{dR}}(Q) \to H^k_{\mathrm{dR}}(M)$;
(2) For the identity map $\mathrm{Id} : M \to M$, $\mathrm{Id}^* : H^k_{\mathrm{dR}}(M) \to H^k_{\mathrm{dR}}(M)$ is the identity map.

It follows in particular that whenever $f : M \to N$ is a diffeomorphism, $f^* : H^k_{\mathrm{dR}}(N) \to H^k_{\mathrm{dR}}(M)$ is a vector space isomorphism for each k.

Proof. The relation $f^*(d\omega) = d(f^*\omega)$ implies that f^* preserves both the spaces of closed forms and exact forms, and thus descends to their quotient. The rest of the statement follows immediately from the basic properties of pullbacks. □

Remark 13.26. For those who enjoy this kind of language, Proposition 13.25 says that H^k_{dR} for each $k \in \mathbb{Z}$ defines a contravariant functor from the category of smooth manifolds and smooth maps to the category of real vector spaces and linear maps.

Example 13.27. The closed 0-forms on M are the locally constant functions, which can take independent but constant values on each connected component of M, while the subspace of exact 0-forms is trivial, thus if M has $N \in \mathbb{N}$ connected components, $H^0_{\mathrm{dR}}(M) \cong \mathbb{R}^N$.

[47] Recall that if $A : V \to W$ is a linear map between vector spaces and $X \subset V$ and $Y \subset W$ are linear subspaces such that $A(X) \subset Y$, then there is a well-defined linear map $V/X \to W/Y$ sending the equivalence class $[x] \in V/X$ of each $x \in V$ to the equivalence class $[Ax] \in W/Y$ of $Ax \in W$. One says in this situation that $A : V \to W$ **descends** to a map $V/X \to W/Y$.

EXAMPLE 13.28. If $M := \{\text{pt}\}$ is the 0-manifold consisting of a single point, then $\Omega^0(\{\text{pt}\}) \cong \mathbb{R}$, $\Omega^k(\{\text{pt}\}) = 0$ for each $k > 0$, and the exterior derivative is the trivial map, implying

$$H^k_{\text{dR}}(\{\text{pt}\}) \cong \begin{cases} \mathbb{R} & \text{for } k = 0, \\ 0 & \text{for } k > 0. \end{cases}$$

EXAMPLE 13.29. Theorem 13.19 implies that $H^1_{\text{dR}}(M) = 0$ whenever M is simply connected.

EXAMPLE 13.30. Corollary 13.9 implies that $H^n_{\text{dR}}(M) \neq 0$ whenever M is a closed oriented n-manifold.

Diffeomorphism-invariance is a nice property, but de Rham cohomology also satisfies a stronger invariance property that makes it much easier to compute.

DEFINITION 13.31. Two smooth maps $f_0, f_1 : M \to N$ are called **smoothly homotopic** (*glatt homotop*) if there exists a smooth map $h : [0,1] \times M \to N$ such that $h(0,\cdot) = f_0$ and $h(1,\cdot) = f_1$.

THEOREM 13.32. *If $f_0, f_1 : M \to N$ are smoothly homotopic maps, then for each k, the linear maps $H^k_{\text{dR}}(N) \to H^k_{\text{dR}}(M)$ defined by f_0^* and f_1^* are identical.*

Before proving this, let's think through some of the consequences. A map $f : M \to N$ is called a **smooth homotopy equivalence** (*glatte Homotopieäquivalenz*) if there exists another smooth map $g : N \to M$ such that $f \circ g : N \to N$ and $g \circ f : M \to M$ are each smoothly homotopic to the identity map. Combining Proposition 13.25 with Theorem 13.32 in this situation implies that $f^* : H^*_{\text{dR}}(N) \to H^*_{\text{dR}}(M)$ and $g^* : H^*_{\text{dR}}(M) \to H^*_{\text{dR}}(N)$ are inverses; in particular, f^* is an isomorphism:

COROLLARY 13.33. *If two manifolds M and N are smoothly homotopy equivalent, then their de Rham cohomologies are isomorphic.* □

The power of Corollary 13.33 lies in the fact that two manifolds can easily be homotopy equivalent without being diffeomorphic; in fact, homotopy equivalence does not even imply that they have the same dimension. Here is an extreme example: a manifold M is called **smoothly contractible** (*glatt zusammenziehbar*) if there exists a smooth homotopy of the identity map $M \to M$ to a constant map. It is easy to see for instance that $\mathbb{R}^n$ is smoothly contractible, and so is any convex subset of $\mathbb{R}^n$. Given a smooth homotopy $h : [0,1] \times M \to M$ with $h(1,\cdot) = \text{Id}_M$ and $h(0,\cdot) \equiv p \in M$ for some fixed point $p \in M$, consider the maps

$$\pi : M \to \{p\}, \qquad i : \{p\} \hookrightarrow M,$$

where π is the unique map and i is the natural inclusion. Now $\pi \circ i$ is the identity map on $\{p\}$, and $i \circ \pi : M \to M$ is $h(0,\cdot)$, which is therefore smoothly homotopic to Id_M. This proves that M is smoothly homotopy equivalent to the one-point manifold $\{p\}$, so combining Corollary 13.33 with Example 13.28 gives:

COROLLARY 13.34. *If M is smoothly contractible, then $H^k_{\text{dR}}(M) = 0$ for all $k > 0$ and $H^0_{\text{dR}}(M) \cong \mathbb{R}$.*

PROOF OF THE POINCARÉ LEMMA. Every point $p \in M$ has a neighborhood $\mathcal{U} \subset M$ that looks like a convex set in some coordinate chart and is thus smoothly contractible. For $k > 0$, it now follows from $H^k_{\text{dR}}(\mathcal{U}) = 0$ that the spaces of closed and exact k-forms on $\mathcal{U}$ are identical. □

PROOF OF THEOREM 13.32. We assume $h : [0,1] \times M \to N$ satisfies $h(0,\cdot) = f_0$ and $h(1,\cdot) = f_1$. Given $\omega \in \Omega^k(N)$, let us assume $L \subset M$ is a compact oriented k-dimensional submanifold with boundary and consider the integral of $h^* d\omega \in \Omega^{k+1}([0,1] \times M)$ over the domain $[0,1] \times L$. Note that the latter is not a smooth manifold with boundary unless $\partial L = \varnothing$; in general $[0,1] \times L$ can be understood as a manifold with boundary *and corners*. Nonetheless, one can make sense of Stokes' theorem on this domain as described in Example 12.14, leading to the relation

$$\begin{aligned}\int_{[0,1]\times L} h^*(d\omega) &= \int_{[0,1]\times L} d(h^*\omega) = \int_{\partial([0,1]\times L)} h^*\omega := \int_{\partial[0,1]\times L} h^*\omega - \int_{[0,1]\times\partial L} h^*\omega \\ &= \int_{\{1\}\times L} h^*\omega - \int_{\{0\}\times L} h^*\omega - \int_{[0,1]\times\partial L} h^*\omega \\ &= \int_L f_1^*\omega - \int_L f_0^*\omega - \int_{[0,1]\times\partial L} h^*\omega,\end{aligned} \tag{13.2}$$

where in the last line we have used the obvious identifications of $\{1\}\times L$ and $\{0\}\times L$ with L, so that the restrictions of $h^*\omega$ to these two submanifolds become $f_1^*\omega$ and $f_0^*\omega$ respectively. Now observe that for any compact oriented m-dimensional submanifold $Q \subset M$ and an $(m+1)$-form $\alpha \in \Omega^{m+1}(N)$, there is a natural way of presenting $\int_{[0,1]\times Q} h^*\alpha$ as the integral of an m-form over Q: we define $P\alpha \in \Omega^m(M)$ namely via the formula

$$(P\alpha)_p(X_1,\dots,X_m) := \int_0^1 (h^*\alpha)_{(t,p)}(\partial_t, X_1,\dots,X_m)\,dt \in \mathbb{R},$$

where ∂_t here denotes the obvious unit vector field on $[0,1]\times M$ pointing in the positive direction on the first factor, and each $X_1,\dots,X_m \in T_pM$ is regarded as living in the subspace $\{0\}\times T_pM \subset T_t[0,1]\times T_pM = T_{(t,p)}([0,1]\times M)$. In this way we have defined a linear operator

$$P : \Omega^{m+1}(N) \to \Omega^m(M) \qquad \text{such that} \qquad \int_{[0,1]\times Q} h^*\alpha = \int_Q P\alpha$$

for all $\alpha \in \Omega^{m+1}(N)$ and compact oriented m-dimensional submanifolds $Q \subset M$. We can use this to transform (13.2) into the relation

$$\int_L (f_1^*\omega - f_0^*\omega) = \int_L P(d\omega) + \int_{\partial L} P\omega = \int_L \left[P(d\omega) + d(P\omega)\right],$$

where we have again applied Stokes' theorem to transform the integral over ∂L into one over L. We now have an equality of the integrals of two k-forms over an arbitrary compact oriented k-dimensional submanifold with boundary: in particular, one could pick any point $p \in M$ and any vectors $X_1,\dots,X_k \in T_pM$ and then approximate the evaluation of both k-forms on $(X_1,\dots,X_k)$ arbitrarily well by integrating them over a submanifold L that is chosen to be a small k-disk through p tangent to the space spanned by $X_1,\dots,X_k$. The conclusion is that these two k-forms must be identical, so we have proved that $f_1^*\omega - f_0^*\omega = P(d\omega) + d(P\omega)$, or rewriting it as an equality between two linear maps $H^k_{\mathrm{dR}}(N) \to H^k_{\mathrm{dR}}(M)$,

$$f_1^* - f_0^* = P\circ d + d\circ P. \tag{13.3}$$

This formula is well known in homological algebra: it is called the **chain homotopy relation**, and the operator $P : \Omega^*(N) \to \Omega^*(M)$ of degree -1 is consequently called a **chain homotopy** (*Kettenhomotopie*). Its existence has the following consequence: if $\omega \in \Omega^k(N)$ is closed, then

$$f_1^*\omega = f_0^*\omega + d(P\omega),$$

implying that $f_1^*\omega$ and $f_0^*\omega$ represent the same element in the quotient $H^k_{\mathrm{dR}}(M)$. □

EXERCISE 13.35. Suppose $\mathcal{O}$ is an open subset of either $\mathbb{H}^n$ or $\mathbb{R}^n$. We call $\mathcal{O}$ a **star-shaped** domain if for every $p \in \mathcal{O}$, it also contains the points $tp \in \mathbb{R}^n$ for all $t \in [0,1]$. It follows that $h(t,p) := tp$ defines a smooth homotopy $h : [0,1]\times\mathcal{O} \to \mathcal{O}$ between the identity and the constant map whose value is the origin, making $\mathcal{O}$ smoothly contractible. Use this homotopy to extract from the proof of Theorem 13.32 an explicit formula for a linear operator $P : \Omega^k(\mathcal{O}) \to \Omega^{k-1}(\mathcal{O})$ for each $k \geqslant 1$ satisfying

$$\omega = P(d\omega) + d(P\omega)$$

for all $\omega \in \Omega^k(\mathcal{O})$. In particular, whenever ω is a closed k-form, $P\omega$ is a primitive of ω. (As a sanity check, a formula for P is given in Remark 13.39 at the end of this lecture, but try to derive it without knowing it in advance.)

One further property of $H^*_{\mathrm{dR}}(M)$ deserves to be mentioned, though a full explanation of it would fall far outside the scope of this course. By a result known as *de Rham's theorem*, $H^k_{\mathrm{dR}}(M)$ is naturally isomorphic to another invariant that is a standard topic in algebraic topology, namely the kth ***singular cohomology*** with real coefficients:

$$H^k_{\mathrm{dR}}(M) \cong H^k(M; \mathbb{R}).$$

The latter is defined for all topological spaces, not just smooth manifolds. As one learns in algebraic topology, $H^k(M; \mathbb{R})$ is often surprisingly easy to compute, and for instance when M is compact, it can be derived from a finite-dimensional chain complex, implying the highly non-obvious fact that

$$\dim H^k_{\mathrm{dR}}(M) < \infty$$

whenever M is compact.

Exercise 13.36. Here is the most basic computation of $H^*_{\mathrm{dR}}(M)$ for a non-contractible manifold: we will show in this exercise that for every $n \in \mathbb{N}$ and $k \in \{0, \dots, n\}$,

$$\dim H^k_{\mathrm{dR}}(S^n) = \begin{cases} 1 & \text{if } k = 0 \text{ or } k = n, \\ 0 & \text{otherwise.} \end{cases} \tag{13.4}$$

Clearly every sphere S^n for $n \geqslant 1$ is connected,[48] so Example 13.27 establishes $H^0_{\mathrm{dR}}(S^n) \cong \mathbb{R}$. For the computation of $H^k_{\mathrm{dR}}(S^n)$ when $k \geqslant 1$, we proceed by induction on n.

(a) Show that if M is a closed oriented n-manifold, then there is a well-defined linear map

$$H^n_{\mathrm{dR}}(M) \to \mathbb{R} : [\omega] \mapsto \int_M \omega, \tag{13.5}$$

and the following conditions are equivalent:
  (i) $H^n_{\mathrm{dR}}(M) \cong \mathbb{R}$;
  (ii) The map (13.5) is an isomorphism;
  (iii) Every $\omega \in \Omega^n(M)$ satisfying $\int_M \omega = 0$ is exact.

(b) Deduce via Exercise 13.16 that (13.4) is correct for $n = 1$.

(c) Suppose M is a closed n-manifold and ω_+, ω_- is a pair of k-forms on $M \times [-1, 1]$ such that $d\omega_+ = d\omega_-$. Show that the following conditions are equivalent:
  (i) $\omega_+ - \omega_-$ is exact;
  (ii) $i_t^*\omega_+ - i_t^*\omega_-$ is an exact k-form on M for every $t \in [-1, 1]$, where $i_t : M \hookrightarrow M \times [-1, 1]$ denotes the inclusion $p \mapsto (p, t)$.
  (iii) There exists a k-form ω on $M \times [-1, 1]$ which matches $\omega_\pm$ near $M \times \{\pm 1\}$ and satisfies $d\omega = d\omega_+ = d\omega_-$.

Hint: First prove the equivalence of (i) and (ii), after convincing yourself that $i_t : M \hookrightarrow M \times [-1, 1]$ is a smooth homotopy equivalence for each t.

[48]The 0-sphere is a discrete set of two points $S^0 = \{1, -1\} \subset \mathbb{R}$, and is thus not connected. That's why we excluded the case $n = 0$ from (13.4).

(d) Under the same assumptions as in part (c), suppose also that M is oriented and $k = n$. Show that the number $\int_{M\times\{t\}} \omega_+ - \int_{M\times\{t\}} \omega_- \in \mathbb{R}$ is the same for any choice of $t \in [-1,1]$.
Hint: Given $-1 \leqslant t_- < t_+ \leqslant 1$, integrate something over $M \times [t_-, t_+]$ and apply Stokes' theorem.

(e) Now given an integer $n \geqslant 2$, assume (13.4) is true for S^{n-1}, and fix $k \in \{1, \dots, n\}$. Regarding S^n as the unit sphere in $\mathbb{R}^{n+1}$ with standard coordinates $(x^1, \dots, x^{n+1})$, we can decompose it into two overlapping n-dimensional disks $S^n = D_+ \cup D_-$ whose intersection looks like $S^{n-1} \times [-1,1]$; specifically, define

$$D_+ := \{x^1 \geqslant -1/2\} \cap S^n, \qquad D_- := \{x^1 \leqslant 1/2\} \cap S^n.$$

Take a moment to convince yourself that there is a diffeomorphism $D_+ \cap D_- \cong S^{n-1} \times [-1,1]$. Observe next that D_+ and D_- are each smoothly contractible, thus any closed k-form ω on S^n will then by exact over each of D_+ and D_-, giving $\alpha_\pm \in \Omega^{k-1}(D_\pm)$ such that $d\alpha_\pm = \omega$ on $D_\pm$. The difficulty is that α_+ and α_- need not match on $D_+ \cap D_-$. Use the inductive hypothesis and the previous steps in this problem to show that if either $1 \leqslant k \leqslant n-1$ or $k = n$ with $\int_{S^n} \omega = 0$, then there exists $\alpha \in \Omega^{k-1}(S^n)$ satisfying $d\alpha = \omega$; show in fact that α can be chosen to match $\alpha_\pm$ on the portions of $D_\pm$ where D_+ and D_- do not overlap. This completes the inductive proof of (13.4).
Hint: The case $k = n$ is trickiest, as you need to use the hypothesis $\int_{S^n} \omega = 0$ to deduce something about α_+ and α_-. What can you say about the integrals of $\alpha_\pm$ over the "equator" $S^{n-1} \cong \{x^1 = 0\} \subset S^n$? Try Stokes' theorem, but be careful with orientations!

EXERCISE 13.37. Show that the wedge product descends to an associative and graded-commutative product $\cup : H^k_{\mathrm{dR}}(M) \times H^\ell_{\mathrm{dR}}(M) \to H^{k+\ell}_{\mathrm{dR}}(M)$, defined by

$$[\alpha] \cup [\beta] := [\alpha \wedge \beta].$$

This is called the **cup product** on de Rham cohomology.
Remark: There is similarly a cup product on singular cohomology, to which this one is isomorphic via de Rham's theorem. But this one is easier to define, and is thus often used in practice as a surrogate for the singular cup product.

EXERCISE 13.38. For this exercise, identify the n-torus $\mathbb{T}^n$ with the quotient $\mathbb{R}^n/\mathbb{Z}^n$ (recall from Exercise 3.4 that there is a natural diffeomorphism). For any sufficiently small open set $\widetilde{\mathcal{U}} \subset \mathbb{R}^n$, the usual Cartesian coordinates $x^1, \dots, x^n : \widetilde{\mathcal{U}} \to \mathbb{R}$ can be used to define a smooth chart $(\mathcal{U}, x)$ on $\mathbb{T}^n$ where

$$\mathcal{U} := \left\{ [p] \in \mathbb{T}^n \;\middle|\; p \in \widetilde{\mathcal{U}} \right\}, \qquad x([p]) := (x^1(p), \dots, x^n(p)) \text{ for } p \in \widetilde{\mathcal{U}}.$$

(a) Show that the coordinate differentials $dx^1, \dots, dx^n \in \Omega^1(\mathcal{U})$ arising from the chart $(\mathcal{U}, x)$ described above are independent of the choice of the set $\widetilde{\mathcal{U}} \subset \mathbb{R}^n$, i.e. the definitions of the coordinate differentials obtained from two different choices $\widetilde{\mathcal{U}}_1, \widetilde{\mathcal{U}}_2 \subset \mathbb{R}^n$ coincide on the region $\mathcal{U}_1 \cap \mathcal{U}_2 \subset \mathbb{T}^n$ where they overlap.

(b) As a consequence of part (a), the 1-forms $dx^1, \dots, dx^n \in \Omega^1(\mathbb{T}^n)$ are well-defined on the entire torus, and they are obviously locally exact and therefore closed, but they might not actually be exact since none of the coordinates $x^1, \dots, x^n$ admit smooth definitions globally on $\mathbb{T}^n$. (This is another example of the phenomenon we saw with $d\theta \in \Omega^1(\mathbb{R}^2 \backslash \{0\})$ in Remark 13.13.) Show in fact that for any vector $(a_1, \dots, a_n) \in \mathbb{R}^n \backslash \{0\}$, the 1-form

$$\lambda := a_i \, dx^i \in \Omega^1(\mathbb{T}^n)$$

is closed but not exact.
Hint: You only need to find one smooth map $\gamma : S^1 \to \mathbb{T}^n$ such that $\int_{S^1} \gamma^* \lambda \neq 0$.

(c) One can similarly produce closed k-forms $\omega \in \Omega^k(\mathbb{T}^n)$ for any $k \leqslant n$ by choosing constants $a_{i_1 \dots i_k} \in \mathbb{R}$ and writing

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} \, dx^{i_1} \wedge \dots \wedge dx^{i_k} \in \Omega^k(\mathbb{T}^n). \tag{13.6}$$

Show that for every nontrivial k-form of this type, one can find a cohomology class $[\alpha] \in H^{n-k}_{\mathrm{dR}}(\mathbb{T}^n)$ such that the cup product $[\omega] \cup [\alpha] \in H^n_{\mathrm{dR}}(\mathbb{T}^n)$ defined in Exercise 13.37 is nontrivial, and deduce from this that ω is not exact.

Hint: Can you choose $\alpha \in \Omega^{n-k}(\mathbb{T}^n)$ so that $\omega \wedge \alpha$ is a volume form?

Remark: One can show that all cohomology classes in $H^k_{\mathrm{dR}}(\mathbb{T}^n)$ are representable by k-forms with constant coefficients as in (13.6), thus $\dim H^k_{\mathrm{dR}}(\mathbb{T}^n) = \binom{n}{k}$.

Remark 13.39. Here is a formula for the operator $P : \Omega^k(\mathcal{O}) \to \Omega^{k-1}(\mathcal{O})$ promised in Exercise 13.35 on a star-shaped domain $\mathcal{O}$ in $\mathbb{H}^n$ or $\mathbb{R}^n$:

$$(P\omega)_p(X_1, \dots, X_{k-1}) := \int_0^1 t^{k-1} \omega_{tp}(p, X_1, \dots, X_{k-1}) \, dt,$$

where since $\mathcal{O}$ is a subset of $\mathbb{R}^n$, we are using the natural isomorphisms $T_p\mathcal{O} = \mathbb{R}^n$ at every point. (Otherwise the expression $\omega_{tp}(p, X_1, \dots, X_{k-1})$ would not generally make sense because $X_1, \dots, X_{k-1} \in T_p\mathcal{O} \neq T_{tp}\mathcal{O}$.) In applications, it is occasionally useful to observe that $P\omega$ depends *continuously* on ω, i.e. one obtains in this way a continuous right-inverse of the operator $d_{k-1} : \Omega^{k-1}(\mathcal{O}) \to \mathrm{im}(d_{k-1}) \subset \Omega^k(\mathcal{O})$.

14. Volume-preserving and symplectic maps

14.1. Volume-preserving flows. Assume M is an oriented n-manifold with a fixed positive volume form $d\mathrm{vol} \in \Omega^n(M)$. In §12.5, we defined the divergence of a vector field $X \in \mathfrak{X}(M)$ in this context as the unique function $\mathrm{div}(X) : M \to \mathbb{R}$ such that

$$d(\iota_X d\mathrm{vol}) = \mathrm{div}(X) \cdot d\mathrm{vol}.$$

A partial justification for this definition was furnished by the Gauss divergence theorem,

$$\int_M \mathrm{div}(X) \, d\mathrm{vol}_M = \int_{\partial M} \langle X, \nu \rangle \, d\mathrm{vol}_{\partial M}, \tag{14.1}$$

a corollary of Stokes' theorem that equates the total divergence of a vector field on a Riemannian manifold with boundary to its total *flux* through the boundary (see §12.5.1). We would now like to explain a more fundamental interpretation of the divergence: it measures the extent to which the flow of X changes volume.

Writing $\mathrm{Vol}(A) := \int_A d\mathrm{vol}$, a diffeomorphism $\varphi : M \to M$ is called **volume preserving** if

$$\mathrm{Vol}(\varphi(A)) = \mathrm{Vol}(A) \qquad \text{for all measurable sets } A \subset M.$$

For a vector field $X \in \mathfrak{X}(M)$ admitting a global flow, we say that its flow is volume preserving if φ^t_X is volume preserving for every $t \in \mathbb{R}$. Without assuming there is a global flow, this condition can still be generalized as follows: for every measurable set $A \subset M$ and every $t \in \mathbb{R}$ for which the domain of φ^t_X contains A, $\mathrm{Vol}(\varphi^t_X(A)) = \mathrm{Vol}(A)$. Note that if A has compact closure, then this condition always makes sense at least for t close to 0. For simplicity we will assume in the following discussion that there is always a global flow, but this condition can be lifted by paying more careful attention to the domains of the flow maps φ^t_X.

The diffeomorphisms $\varphi^t_X : M \to M$ defined via the flow of a vector field are always orientation preserving—this results from the fact that $\varphi^0_X : M \to M$ is the identity map, so for any $p \in M$, any positively oriented basis $Y_1, \dots, Y_n$ of T_pM gives rise to a continuous 1-parameter family of bases

$$(T\varphi_X^t(Y_1), \dots, T\varphi_X^t(Y_n))$$

for the tangent spaces $T_{\varphi_X^t(p)}M$, and continuity dictates that they must all be positively oriented. We therefore have

$$\mathrm{Vol}(\varphi_X^t(A)) = \int_{\varphi_X^t(A)} d\mathrm{vol} = \int_A (\varphi_X^t)^* d\mathrm{vol}$$

for every $A \subset M$, and the rate of change of this volume is

$$\frac{d}{dt}\mathrm{Vol}(\varphi_X^t(A)) = \frac{d}{dt}\int_A (\varphi_X^t)^* d\mathrm{vol} = \int_A \partial_t(\varphi_X^t)^* d\mathrm{vol}. \tag{14.2}$$

The next step in the calculation works in more general contexts: in place of the volume form $d\mathrm{vol}$, we can consider an arbitrary tensor field $S \in \Gamma(T_\ell^k M)$. Recall that $\varphi_X^{s+t} = \varphi_X^s \circ \varphi_X^t$, thus $(\varphi_X^{s+t})^* = (\varphi_X^t)^*(\varphi_X^s)^*$, and

$$\begin{aligned} \partial_t(\varphi_X^t)^* S &= \partial_s(\varphi_X^{s+t})^* S\big|_{s=0} = \partial_s(\varphi_X^t)^*(\varphi_X^s)^* S\big|_{s=0} \\ &= (\varphi_X^t)^*\left(\partial_s(\varphi_X^s)^* S|_{s=0}\right) = (\varphi_X^t)^*\left(\mathcal{L}_X S\right). \end{aligned} \tag{14.3}$$

Applying this to (14.2) gives

$$\frac{d}{dt}\mathrm{Vol}(\varphi_X^t(A)) = \int_A (\varphi_X^t)^*\left(\mathcal{L}_X d\mathrm{vol}\right) = \int_{\varphi_X^t(A)} \mathcal{L}_X d\mathrm{vol}.$$

It follows that the flow is volume preserving if the Lie derivative of the volume form $d\mathrm{vol}$ with respect to X vanishes, and conversely, the derivative of $\mathrm{Vol}(\varphi_X^t(A))$ can only vanish for every measurable set $A \subset M$ if the n-form $(\varphi_X^t)^*(\mathcal{L}_X d\mathrm{vol})$ vanishes identically for every t, which is equivalent to the condition $\mathcal{L}_X d\mathrm{vol} \equiv 0$ since $(\varphi_X^t)^* : \Omega^n(M) \to \Omega^n(M)$ is a bijection.

Lemma 14.1. *For any volume form* $d\mathrm{vol} \in \Omega^n(M)$ *and vector field* $X \in \mathfrak{X}(M)$*,*

$$\mathcal{L}_X d\mathrm{vol} = d(\iota_X d\mathrm{vol}).$$

This relation will follow from the more general formula of Cartan for Lie derivatives of differential forms, to be proved in the next section. We can now alternatively characterize the divergence of X as the unique function such that

$$\mathcal{L}_X d\mathrm{vol} = \mathrm{div}(X) \cdot d\mathrm{vol}, \tag{14.4}$$

and the discussion above implies:

Theorem 14.2. *On a manifold M with volume form* $d\mathrm{vol}$*, a vector field $X \in \mathfrak{X}(M)$ has a volume-preserving flow if and only if* $\mathrm{div}(X) \equiv 0$. □

The divergence theorem (14.1) now admits a new geometric interpretation whenever M is a compact submanifold with boundary in a larger n-manifold N on which the vector field X and volume form $d\mathrm{vol}$ are defined. In this case, the flow φ_X^t of X is well defined on M for all t sufficiently close to zero, and the left hand side of (14.1) then becomes

$$\begin{aligned} \int_M \mathrm{div}(X)\, d\mathrm{vol}_N &= \int_M \mathcal{L}_X(d\mathrm{vol}_N) = \frac{d}{dt}\int_M (\varphi_X^t)^* d\mathrm{vol}_N\bigg|_{t=0} = \frac{d}{dt}\int_{\varphi_X^t(M)} d\mathrm{vol}_N\bigg|_{t=0} \\ &= \frac{d}{dt}\mathrm{Vol}(\varphi_X^t(M))\bigg|_{t=0}. \end{aligned}$$

The divergence theorem thus relates the rate of change of the volume of M under the flow of X to the average of $\langle X, \nu\rangle$ along ∂M, which measures the extent to which X flows out of M vs. into M through its boundary.

14.2. Cartan's formula for the Lie derivative. The following practical tool for computing Lie derivatives of forms is sometimes called *Cartan's magic formula.*

THEOREM 14.3. *For any $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^k(M)$,*

$$\mathcal{L}_X \omega = d(\iota_X \omega) + \iota_X(d\omega).$$

An immediate application is Lemma 14.1 above: if $d\mathrm{vol} \in \Omega^n(M)$ is a volume form, then

$$\mathcal{L}_X d\mathrm{vol} = d(\iota_X d\mathrm{vol}) + \iota_X d(d\mathrm{vol}) = d(\iota_X d\mathrm{vol})$$

since $d(d\mathrm{vol})$ is an $(n+1)$-form on an n-manifold and therefore vanishes.[49]

The following sequence of exercises sums up to a proof of Cartan's formula, the idea behind it being to show that for any given $X \in \mathfrak{X}(M)$, both of the operators $\mathcal{L}_X$ and $d\iota_X + \iota_X d$ define derivations on the exterior algebra $\Omega^*(M)$ that match when applied to functions or differentials of functions. This is sufficient for the same reason that a few formal properties centered around the graded Leibniz rule sufficed in Proposition 9.16 for characterizing the exterior derivaive: both are clearly local operators, and locally, every differential form is a finite sum of wedge products of functions and differentials.

EXERCISE 14.4 (easy). Show that Theorem 14.3 holds for all $\omega = f \in C^\infty(M) = \Omega^0(M)$.

LEMMA 14.5. *Theorem 14.3 holds for all $\omega = df \in \Omega^1(M)$ with $f \in C^\infty(M)$.*

PROOF. Since $d^2 = 0$, $d\iota_X df + \iota_X d(df) = d(\iota_X df)$, where $\iota_X df$ is the real-valued function $p \mapsto df(X(p))$. To evaluate $\mathcal{L}_X(df) \in \Omega^1(M)$ on some $Y \in T_pM$ at a point $p \in M$, choose a smooth path $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ and $\dot\gamma(0) = Y$. Then using Proposition 9.18,

$$\begin{aligned}\mathcal{L}_X(df)(Y) &= \partial_t(\varphi_X^t)^*(df)(Y)\big|_{t=0} = \partial_t d(f \circ \varphi_X^t)(Y)\big|_{t=0} = \partial_t \partial_s f(\varphi_X^t(\gamma(s)))\big|_{s=t=0} \\ &= \partial_s \partial_t f(\varphi_X^t(\gamma(s)))\big|_{s=t=0} = \partial_s df(X(\gamma(s))\big|_{s=0} = \partial_s \iota_X(df)(\gamma(s))\big|_{s=0} = d(\iota_X df)(Y).\end{aligned}$$

□

The next exercise follows also quite easily from the definition of the Lie derivative, plus Proposition 9.18 and the fact that the wedge product is bilinear. Notice that in contrast to the exterior derivative, no annoying sign appears in the Leibniz rule for $\mathcal{L}_X$. Formally, the reason is because $\mathcal{L}_X$ sends k-forms to k-forms for each $k \geqslant 0$, and is thus an operator of "degree 0", i.e. it is even, while the exterior derivative is odd.

EXERCISE 14.6. Show that $\mathcal{L}_X : \Omega^*(M) \to \Omega^*(M)$ is a derivation with respect to the wedge product, meaning

$$\mathcal{L}_X(\alpha \wedge \beta) = \mathcal{L}_X \alpha \wedge \beta + \alpha \wedge \mathcal{L}_X \beta.$$

We now turn our attention fully to the operator

$$P_X := d\iota_X + \iota_X d : \Omega^*(M) \to \Omega^*(M), \tag{14.5}$$

in which each term is a composition of operators with degrees 1 and -1, so P_X itself also has degree 0. We've seen already that d satisfies a graded Leibniz rule; it turns out that ι_X does as well:

[49] Here is another cautionary reminder about the oddity of our notation for volume forms: we have not defined any $(n-1)$-form "$\mathrm{vol} \in \Omega^{n-1}(M)$" for $d\mathrm{vol}$ to be the exterior derivative of, and we have seen for instance that when M is a closed manifold, $d\mathrm{vol}$ is definitely not the exterior derivative of anything. The vanishing of $d(d\mathrm{vol})$ thus has nothing to do with the relation $d \circ d = 0$; it vanishes for a completely different reason.

EXERCISE 14.7. For V an n-dimensional vector space, the goal of this exercise is to show that for every $v \in V$, the operator $\iota_v : \Lambda^* V^* \to \Lambda^* V^*$ satisfies the graded Leibniz rule

$$\iota_v(\alpha \wedge \beta) = (\iota_v \alpha) \wedge \beta + (-1)^k \alpha \wedge (\iota_v \beta) \tag{14.6}$$

for all $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^\ell V^*$. The statement is trivial if $v = 0$, so assume otherwise, in which case we may as well assume v is the first element e_1 of a basis $e_1, \ldots, e_n \in V$, whose dual basis we can denote by $e_*^1, \ldots, e_*^n \in V^* = \Lambda^1 V^*$.

(a) Prove that (14.6) holds whenever α and β are both products of the form $\alpha = e_*^{i_1} \wedge \ldots \wedge e_*^{i_k}$ and $\beta = e_*^{j_1} \wedge \ldots \wedge e_*^{j_\ell}$ with $i_1 < \ldots < i_k$ and $j_1 < \ldots < j_\ell$.
Hint: Consider separately a short list of cases depending on whether each of i_1 and j_1 are 1 and whether the sets $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_\ell\}$ are disjoint.

(b) Deduce via linearity that (14.6) holds always.

EXERCISE 14.8. Prove that the operator P_X in (14.5) is also a derivation on $\Omega^*(M)$, and deduce that $P_X = \mathcal{L}_X$, thus proving Theorem 14.3.

14.3. Symplectic manifolds and Hamiltonian systems. Volume-preserving flows arise naturally in the context of Hamiltonian systems, a special class of dynamical systems that originate in classical mechanics. From a mathematical perspective, the most natural language for this discussion is that of *symplectic* geometry.

DEFINITION 14.9. Assume M is a smooth manifold of even dimension $2n$ for some $n \in \mathbb{N}$. A 2-form $\omega \in \Omega^2(M)$ is called **symplectic** (*symplektisch*) if every point $x \in M$ admits a neighborhood $x \in \mathcal{U} \subset M$ with a coordinate chart of the form $(\mathcal{U}, (p^1, q^1, \ldots, p^n, q^n))$ such that

$$\omega = \sum_{j=1}^n dp^j \wedge dq^j \qquad \text{on } \mathcal{U}. \tag{14.7}$$

A 2-form with this property is also sometimes called a **symplectic structure** (*symplektische Struktur*) on M, and the pair (M, ω) in this situation is called a **symplectic manifold** (*symplektische Mannigfaltigkeit*).

Observe that the coordinates $(p^1, q^1, \ldots, p^n, q^n)$ appearing in (14.7) are special; it would certainly be impossible to demand that any 2-form satisfy (14.7) for *every* choice of chart, but the definition only requires the existence of *some* chart near every point so that ω takes this form. In this sense, a symplectic structure is somewhat analogous to an orientation: it is equivalent in fact to a maximal atlas of compatible charts in which the word "compatible" has been given a new and much stricter definition, requiring all transition maps to not only be smooth but also to preserve the relation (14.7). Physicists sometimes refer to coordinates $(p^1, q^1, \ldots, p^n, q^n)$ of this type as *canonical coordinates* and call the corresponding transition maps *canonical transformations*. Mathematicians prefer to call them *Darboux coordinates*, after Darboux's theorem (see Remark 14.11 below).

EXERCISE 14.10. Show that a symplectic form $\omega \in \Omega^2(M)$ always has the following properties:

(a) ω is closed: $d\omega = 0$.

(b) For every $x \in M$, the linear map $T_x M \to T_x^* M : X \mapsto \omega(X, \cdot)$ is an isomorphism. (Bilinear forms with this property are called **nondegenerate**).

(c) The "top" exterior power of ω,

$$\omega^n := \underbrace{\omega \wedge \ldots \wedge \omega}_{n} \in \Omega^{2n}(M)$$

is a volume form on M. It follows in particular that M is orientable.

(d) If M is closed, then ω represents a nontrivial cohomology class $[\omega] \in H^2_{\mathrm{dR}}(M)$.
Hint: Recall the cup product from Exercise 13.37. What can you say about the n-fold cup product of $[\omega]$ with itself?

Remark 14.11. A fundamental result known as *Darboux's theorem* says that symplectic forms can in fact be characterized fully in terms of the first two properties in Exercise 14.10, i.e. every 2-form that is both closed and nondegenerate admits an atlas of charts satisfying (14.7) and is thus a symplectic form. This reveals for instance that every volume form on a surface[50] is a symplectic form. We will not make use of these facts here, but it is important to be aware of them since most textbooks prefer to *define* the term "symplectic form" to mean a closed and nondegenerate 2-form.

Given a smooth function $H : M \to \mathbb{R}$ on a symplectic manifold (M, ω), the nondegeneracy of ω implies that there is a unique vector field $X_H \in \mathfrak{X}(M)$ satisfying

$$\omega(X_H, \cdot) = -dH \in \Omega^1(M). \tag{14.8}$$

We call X_H the **Hamiltonian vector field** determined by H, and in this context, the function H itself is often called a **Hamiltonian**. In Darboux coordinates, it is not hard to derive an explicit formula for the Hamiltonian vector field: writing $X_H = A^j \frac{\partial}{\partial q^j} + B^j \frac{\partial}{\partial p^j}$, we find

$$\begin{aligned} dH = \frac{\partial H}{\partial q^j}\, dq^j + \frac{\partial H}{\partial p^j}\, dp^j &= -\omega(X_H, \cdot) = -\sum_{i=1}^n (dp^i \wedge dq^i) \left(A^j \frac{\partial}{\partial q^j} + B^j \frac{\partial}{\partial p^j}, \cdot \right) \\ &= \sum_{i=1}^n \left(-B^i\, dq^i + A^i\, dp^i \right), \end{aligned}$$

implying

$$X_H = \sum_{i=1}^n \left(\frac{\partial H}{\partial p^i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p^i} \right). \tag{14.9}$$

In other words, if $x(t) \in M$ denotes a smooth path passing through the domain of a Darboux chart and its coordinates in this chart at time t are written as $(p^1(t), q^1(t), \dots, p^n(t), q^n(t))$, then x is an orbit of X_H if and only if its coordinates satisfy the following system of $2n$ first-order ODEs:

$$\dot{q}^i(t) = \frac{\partial H}{\partial p^i}(x(t)), \qquad \dot{p}^i(t) = -\frac{\partial H}{\partial q^i}(x(t)) \qquad i = 1, \dots, n. \tag{14.10}$$

This system is known as *Hamilton's equations*, and the dynamical system defined by the flow of X_H is called a *Hamiltonian system*.

The study of Hamiltonian systems originates with the following example.

Example 14.12. In classical mechanics, the motion in $\mathbb{R}^3$ of a single particle with mass $m > 0$ under the influence of a force is described by a path $\mathbf{q}(t) = (q^1(t), q^2(t), q^3(t)) \in \mathbb{R}^3$ that obeys Newton's second law,

$$\mathbf{F}(\mathbf{q}(t)) = m\ddot{\mathbf{q}}(t),$$

where $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field representing the force. In standard examples, $\mathbf{F}$ is determined by a *potential* $V : \mathbb{R}^3 \to \mathbb{R}$ via the relation

$$\mathbf{F} = -\nabla V,$$

[50] On a manifold of dimension 2, it is also common to refer to volume forms as **area forms**.

hence the individual coordinates satisfy $m\ddot{q}^i(t) = -\frac{\partial V}{\partial q^i}(\mathbf{q}(t))$. There is a popular trick for turning second-order systems of ODEs like this one into first-order systems with twice as many degrees of freedom: we associate to the position variables q^1, q^2, q^3 the corresponding *momentum* variables

$$p^i(t) := m\dot{q}^i(t), \qquad \mathbf{p} := (p^1, p^2, p^3)$$

and observe that the path $(\mathbf{q}(t), \mathbf{p}(t)) \in \mathbb{R}^6$ now satisfies the first-order system of equations

$$\dot{q}^i(t) = \frac{1}{m}p^i(t), \qquad \dot{p}^i(t) = -\frac{\partial V}{\partial q^i}(\mathbf{q}(t)), \qquad i = 1, 2, 3.$$

As it happens, this is the Hamiltonian system determined by the function $H : \mathbb{R}^6 \to \mathbb{R}$ given by

$$H(\mathbf{q}, \mathbf{p}) := \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{q}).$$

Rewriting this as a function of $\mathbf{q}$ and $\dot{\mathbf{q}} := \frac{1}{m}\mathbf{p}$, the first term becomes $\frac{1}{2}m|\dot{\mathbf{q}}|^2$, which physicists call the *kinetic energy* of the moving particle. This is summed with the potential energy $V(\mathbf{q})$ to produce the Hamiltonian, which therefore has an interpretation as the *total energy* of the particle.

The Hamiltonian formalism lends itself to generalization: to turn the example above into a system of $N > 1$ moving particles, one can package the coordinates of all particles together to form a path in $\mathbb{R}^{3N}$, define corresponding momenta to produce a path in the so-called **phase space** $\mathbb{R}^{6N}$, write the total energy of the system as a function of all its position and momentum variables, and then write down Hamilton's equations (14.10). More generally, one can consider systems with constraints that prevent their positions from moving freely in Euclidean space, but confine them instead to a submanifold. In this situation there might not exist any global coordinate system in which Hamilton's equations (14.10) make sense, but if we have a symplectic form and a Hamiltonian function, then (14.8) defines the Hamiltonian vector field in a way that is independent of coordinates. We will see for instance that on any n-dimensional Riemannian manifold, the geodesic equation can be identified with a Hamiltonian system on a manifold of dimension $2n$.

If you've wondered why we are discussing symplectic manifolds in the same lecture with volume-preserving flows, here is the reason:

THEOREM 14.13 (Liouville's theorem). *For any symplectic manifold (M, ω) and Hamiltonian $H \in C^\infty(M)$, the flow of the Hamiltonian vector field X_H is volume preserving with respect to the volume form $\omega^n \in \Omega^{2n}(M)$.*

PROOF. Let's do two proofs. The first is a coordinate-based computation: in any Darboux chart on some region in M, ω^n becomes a constant multiple of the standard volume form

$$\omega^n = \left(\sum_{i_1=1}^n dp^{i_1} \wedge dq^{i_1}\right) \wedge \ldots \wedge \left(\sum_{i_n=1}^n dp^{i_n} \wedge dq^{i_n}\right) = n\, dp^1 \wedge dq^1 \wedge \ldots \wedge dp^n \wedge dq^n,$$

and according to Exercise 12.16 and (14.9), the divergence of X_H is thus

$$\operatorname{div}(X_H) = \sum_{i=1}^n \left(\frac{\partial}{\partial q^i}\frac{\partial H}{\partial p^i} - \frac{\partial}{\partial p^i}\frac{\partial H}{\partial q^i}\right) = 0.$$

The result now follows from Theorem 14.2.

The second proof is more elegant, because it does not require coordinates, and it also proves a stronger result. Using Cartan's formula and the defining property of the vector field X_H, we find

$$\mathcal{L}_{X_H}\omega = d(\iota_{X_H}\omega) + \iota_{X_H}(d\omega) = -d(dH) = 0.$$

It follows via (14.3) that the 2-forms $(\varphi^t_{X_H})^*\omega$ are independent of t, and thus

$$(\varphi^t_{X_H})^*\omega = \omega \text{ for all } t. \tag{14.11}$$

It follows that for each t, $\varphi := \varphi^t_{X_H}$ also preserves the volume form ω^n, since

$$\varphi^*(\omega \wedge \ldots \wedge \omega) = \varphi^*\omega \wedge \ldots \wedge \varphi^*\omega = \omega \wedge \ldots \wedge \omega. \tag{14.12}$$

□

I mentioned that our second proof of Liouville's theorem actually proves a stronger result. On a symplectic manifold (M, ω), a diffeomorphism $\psi : M \to M$ that satisfies

$$\psi^*\omega = \omega$$

is called a **symplectomorphism** (*Symplektomorphismus*), which can be viewed as an abbreviation for **symplectic diffeomorphism**. We see from (14.11) that Hamiltonian flows $\varphi^t_{X_H}$ have this property for every t, and by (14.12), all symplectomorphisms are also volume preserving.

While the subject of symplectic geometry has existed since the beginning of the 20th century, it was unknown for many decades whether the condition of being a symplectomorphism is truly more restrictive than being volume preserving. The following answer to this question emerged in 1985 and opened up a whole new subfield of geometry, known as *symplectic topology*:

Theorem (Gromov's non-squeezing theorem [Gro85]). *Fix the global coordinates $(p^1, q^1, \ldots, p^n, q^n)$ on $\mathbb{R}^{2n}$ with the "standard" symplectic form $\omega = \sum_{i=1}^n dp^i \wedge dq^i$, and let $B^k_r \subset \mathbb{R}^k$ denote the open ball of radius r. Then for two constants $r, R > 0$, the $2n$-ball $B^{2n}_r \subset \mathbb{R}^{2n}$ is symplectomorphic to a subset of the "cylinder"*

$$Z^{2n}_R := B^2_R \times \mathbb{R}^{2n-2} \subset \mathbb{R}^{2n}$$

if and only if $r \leqslant R$.

This is a hard theorem; various proofs are known, but all of them require a substantial amount of analytical machinery which cannot be fit into an introductory course. The significance of the non-squeezing theorem is that if $n \geqslant 2$, then no matter how small $R > 0$ may be, the cylinder Z^{2n}_R contains unlimited space in $2n - 2$ of its $2n$ dimensions, and it is never difficult to find a volume-preserving embedding $B^{2n}_r \hookrightarrow Z^{2n}_R$ that compresses the first two dimensions as much as needed while expanding the others to compensate. The fact that *symplectic* embeddings cannot do this when $R < r$ means that there are meaningful restrictions on symplectic maps beyond the requirement that they must preserve volume. That subject is still an active area of research today.

Exercise 14.14. In 1915, Emmy Noether established a beautiful correspondence between the conserved quantities of a mechanical system and its symmetries. A simple version of this theorem in the Hamiltonian context takes the following form. Assume (M, ω) is a symplectic manifold, and $H : M \to \mathbb{R}$ and $F : M \to \mathbb{R}$ are two functions such that the corresponding Hamiltonian vector fields X_H and X_F have global flows. We say that F is *conserved* under the flow of X_H if F is constant along every orbit of X_H.

(a) Show that F is conversed under the flow of X_H if and only if H is conserved under the flow of X_F.

(b) In some settings, there is a converse to the result proved in part (a). Suppose M is simply connected, and $Y \in \mathfrak{X}(M)$ is a vector field with a global flow that is symplectic and preserves H, i.e.

$$(\varphi^t_Y)^*\omega = \omega \qquad \text{and} \qquad H \circ \varphi^t_Y = H \tag{14.13}$$

for all t. One says in this situation that Y determines a *symmetry* of the Hamiltonian system on (M, ω) defined by H. Under these assumptions, show that there exists a function $F : M \to \mathbb{R}$, uniquely defined up to addition of a constant, such that $Y = X_F$, and F is then conserved under the flow of X_H.

Let's work out a concrete example. Let $M = \mathbb{R}^4$ with coordinates (p_x, x, p_y, y) and the standard symplectic form

$$\omega = dp_x \wedge dx + dp_y \wedge dy \in \Omega^2(\mathbb{R}^4).$$

We can think of $\mathbb{R}^4$ as the "position-momentum space" (also known as *phase space*) representing the motion of a single particle of mass $m > 0$ in a plane: its position is given by $\mathbf{q} := (x, y) \in \mathbb{R}^2$, and $\mathbf{p} := (p_x, p_y) \in \mathbb{R}^2$ are the corresponding momentum variables. Given a "potential" function $V : \mathbb{R}^2 \to \mathbb{R}$, the total energy of the system is given by the function $H : \mathbb{R}^4 \to \mathbb{R}$,

$$H = \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{q}).$$

Suppose now that the potential V is chosen to be *rotationally symmetric*, e.g. this is the case if $\mathbf{q}$ represents the position of the Earth moving around the sun (with the latter positioned at the origin). To express this condition succinctly, one can transform to polar coordinates (r, θ) on a suitable subset of $\mathbb{R}^2$, related to the (x, y)-coordinates as usual by $x = r\cos\theta$ and $y = r\sin\theta$. The condition imposed on V is then $\partial_\theta V \equiv 0$.

(c) Regarding r and θ as real-valued functions on (a suitable subdomain of) $\mathbb{R}^4$ that depend on the coordinates x and y but not on p_x and p_y, define two additional functions on the same domain by

$$p_r := \frac{x}{r}p_x + \frac{y}{r}p_y, \qquad p_\theta := yp_x - xp_y.$$

Show that $(p_r, r, p_\theta, \theta)$ is then a Darboux chart for the symplectic form ω.
Hint: It suffices to compute ω in the new coordinates and show that it satisfies the right formula, but this computation is a bit long. You could make your life easier by observing that $\omega = d\lambda$ for $\lambda := p_x\, dx + p_y\, dy$, and then computing λ in the new coordinates.

(d) Write down H as a function of $(p_r, r, p_\theta, \theta)$ and show that the vector field $Y := \partial_\theta$ defined in these coordinates on $\mathbb{R}^4 \backslash \{r = 0\}$ satisfies (14.13). Derive a formula for the corresponding conserved quantity F as promised by part (b). It is called the angular momentum of the system.

15. Partitions of unity

In Lecture 11, we constructed partitions of unity subordinate to finite open covers of compact manifolds: more precisely, if $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ is a finite collection of open sets in a manifold M whose union contains the compact subset $K \subset M$, then there exists an associated collection of smooth functions $\{\varphi_\alpha : M \to [0, 1]\}_{\alpha \in I}$ such that

$$\sum_{\alpha \in I} \varphi_\alpha \equiv 1 \text{ on } K, \qquad \text{and} \qquad \operatorname{supp}(\varphi_\alpha) \subset \mathcal{U}_\alpha \text{ is compact for every } \alpha \in I.$$

This was used in order to "localize" the problem of defining integrals $\int_A \omega$, and we used the same localization trick again to prove Stokes' theorem in Lecture 12. In this lecture, we will use a more general localization trick to prove that Riemannian metrics exist on all smooth manifolds M. Unless M happens to be compact, we will not be able to get away with considering only finite open covers or functions with compact support. We will therefore need a more general notion of partitions of unity and an extension of the previous construction. This turns out to be the point where one must finally make explicit use of the assumption that manifolds are metrizable.

15.1. Local finiteness. A collection of subsets $\{\mathcal{U}_\alpha \subset X\}_{\alpha \in I}$ in a topological space X is called **locally finite** if every point $p \in X$ has a neighborhood that intersects at most finitely many of the sets $\mathcal{U}_\alpha$. Similarly, a collection of functions $\{f_\alpha : X \to \mathbb{R}\}_{\alpha \in I}$ is called **locally finite** if the sets $\{f_\alpha^{-1}(\mathbb{R}\backslash\{0\}) \subset X\}_{\alpha \in I}$ form a locally finite collection. This condition has the following advantage: if $\{f_\alpha : M \to \mathbb{R}\}_{\alpha \in I}$ is a locally finite collection of *smooth* functions on a manifold M, then one can make sense of the sum

$$\sum_{\alpha \in I} f_\alpha(p) \in \mathbb{R}$$

for every $p \in M$ since, even if I is an uncountably infinite set, at most finitely many terms in this sum are nonzero. Even better, p admits a neighborhood $\mathcal{V} \subset M$ that intersects at most finitely many of the sets $f_\alpha^{-1}(\mathbb{R}\backslash\{0\})$, implying that at most finitely many of the functions f_α can have nonzero values anywhere on $\mathcal{V}$, and $\sum_{\alpha \in I} f_\alpha$ therefore makes sense as a *smooth* function on $\mathcal{V}$. We therefore obtain a global smooth function

$$\sum_{\alpha \in I} f_\alpha \in C^\infty(M),$$

even if the sum contains uncountably many terms that are (somewhere) nontrivial functions on M.

Exercise 15.1. Show that if X is a topological space with open subset $\mathcal{U} \subset X$ and a locally finite collection of continuous functions $\{f_\alpha : X \to \mathbb{R}\}_{\alpha \in I}$ satisfying $\operatorname{supp}(f_\alpha) \subset \mathcal{U}$ for every $\alpha \in \mathcal{U}$, then $\sum_{\alpha \in I} f_\alpha$ also has support in $\mathcal{U}$.

Definition 15.2. Given an open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ of a smooth manifold M, a **partition of unity** subordinate to $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ is a locally finite collection of smooth functions $\{\varphi_\alpha : M \to [0,1]\}_{\alpha \in I}$ which satisfy the following assumptions:

(1) For each $\alpha \in I$, $\operatorname{supp}(\varphi_\alpha) \subset \mathcal{U}_\alpha$;
(2) $\sum_{\alpha \in I} \varphi_\alpha \equiv 1$.

Note that in Definition 15.2, the condition $\sum_{\alpha \in I} \varphi_\alpha \equiv 1$ makes sense due to the local finiteness assumption; this condition was automatic in Lecture 11 because we were considering only a finite collection of functions, but here we are not assuming the collection is finite, nor that the functions have compact support. This relaxation of assumptions makes it possible to prove the following without assuming M is compact:

Theorem 15.3. *Every open cover of a smooth manifold admits a subordinate partition of unity.*

This theorem will be proved in §15.4.

15.2. Existence of Riemannian metrics and volume forms. Before proving that partitions of unity always exist, we shall demonstrate their usefulness by proving the following:

Theorem 15.4. *Every smooth manifold admits a Riemannian metric.*

As a preliminary remark relevant to the proof, we observe that on any vector space V, the set of inner products on V forms a *convex* subset of the vector space of bilinear maps $V \times V \to \mathbb{R}$. Indeed, the symmetric bilinear maps form a linear subspace, and whenever $\langle\ ,\ \rangle_0$ and $\langle\ ,\ \rangle_1$ are two inner products on V, the interpolation $\langle\ ,\ \rangle_t := t\langle\ ,\ \rangle_1 + (1-t)\langle\ ,\ \rangle_0$ for $t \in [0,1]$ also satisfies

$$\langle v, v\rangle_t = t\langle v, v\rangle_1 + (1-t)\langle v, v\rangle_0 > 0$$

for every nonzero $v \in V$. More generally, any *convex combination* of finitely many inner products on V is also an inner product, i.e. for any finite collection of inner products $\langle\ ,\ \rangle_i$ and numbers

$\tau_i \in [0,1]$ for $i = 1, \ldots, k$ with $\sum_{i=1}^k \tau_i = 1$,

$$\sum_{i=1}^k \tau_i \langle \ , \ \rangle_i$$

is an inner product.

LEMMA 15.5. *Suppose $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ is an open cover of a smooth manifold M with subordinate partition of unity $\{\varphi_\alpha\}_{\alpha \in I}$, and for each $\alpha \in I$, $g_\alpha \in \Gamma(T_2^0 \mathcal{U}_\alpha)$ is a Riemannian metric on the open subset $\mathcal{U}_\alpha$. Then the formula*

$$g := \sum_{\alpha \in I} \varphi_\alpha g_\alpha$$

defines a Riemannian metric on M, where in this sum, the term $\varphi_\alpha g_\alpha$ is interpreted as an element of $\Gamma(T_2^0 M)$ that vanishes outside of $\mathcal{U}_\alpha$.

PROOF. Since $\mathrm{supp}(\varphi_\alpha) \subset \mathcal{U}_\alpha$, the tensor field $\varphi_\alpha g_\alpha \in \Gamma(T_2^0 \mathcal{U}_\alpha)$ can be extended to a smooth tensor field on M that vanishes outside of $\mathcal{U}_\alpha$, and we will continue to denote the extension by $\varphi_\alpha g_\alpha \in \Gamma(T_2^0 M)$. The sum then makes sense and is smooth due to local finiteness, as every point is contained in a neighborhood on which only finitely many terms of the sum are nontrivial. Moreover, at each individual point $p \in M$, $g_p : T_pM \times T_pM \to \mathbb{R}$ is a convex combination of inner products, and is therefore also an inner product. □

PROOF OF THEOREM 15.4. Choose an open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ of M such that each $\mathcal{U}_\alpha$ is the domain of a chart x_α, and define a Riemannian metric g_α on $\mathcal{U}_\alpha$ that looks like the standard Euclidean inner product in the chosen coordinates. A global Riemannian metric $g \in \Gamma(T_2^0 M)$ can then be defined via Lemma 15.5 after choosing a subordinate partition of unity. □

In light of Corollary 11.10 on the Riemannian volume form associated to a Riemannian metric, Theorem 15.4 implies:

COROLLARY 15.6. *Every smooth oriented manifold admits a volume form.* □

EXERCISE 15.7. Use a partition of unity to prove Corollary 15.6 without mentioning Theorem 15.4 or Riemannian metrics. Use instead the fact that for any oriented n-dimensional vector space V, the set

$$\left\{\omega \in \Lambda^n V^* \;\middle|\; \omega(v_1, \ldots, v_n) > 0 \text{ for some positively-oriented basis } v_1, \ldots, v_n \in V\right\}$$

is convex.

REMARK 15.8. Without assuming M is oriented, Theorem 15.4 also implies that every smooth manifold admits a volume element (see §11.4).

15.3. Paracompactness. Any Riemannian manifold (M, g) is also a metric space in a natural way, at least if it is connected, because one can define the distance between two points $p, q \in M$ by

$$\mathrm{dist}(p,q) := \inf_\gamma \int_a^b \sqrt{g(\dot\gamma(t), \dot\gamma(t))}\, dt, \tag{15.1}$$

where the infimum is over all intervals $[a,b] \subset \mathbb{R}$ and smooth paths $\gamma : [a,b] \to M$ with $\gamma(a) = p$ and $\gamma(b) = q$. For a Riemannian manifold with multiple connected components, each component has a natural metric defined in this way, and there are standard tricks for defining metrics on any disjoint union of metric spaces (see e.g. Exercise 2.23). The point is: if you hadn't already assumed that smooth manifolds are metrizable but you assumed that Theorem 15.4 is true, then the theorem would imply metrizability.

EXERCISE 15.9. Take a moment to convince yourself that (15.1) really does define a metric, in particular that it satisfies the triangle inequality.
Hint: One can reparametrize the path $\gamma : [a, b] \to M$ quite freely without changing the integral. If you take advantage of this freedom, then a path from p to q and a path from q to r can always be concatenated smoothly.

The existence of the metric (15.1) is a dead giveaway that something about Theorem 15.4 depends on our assumption that all manifolds are metrizable. We haven't used that assumption in this course until now. But we will need it for constructing the partition of unity.

Recall that a **refinement** of an open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ is another open cover $\{\mathcal{V}_\beta\}_{\beta \in J}$ such that for every $\beta \in J$, $\mathcal{V}_\beta$ is contained in $\mathcal{U}_\alpha$ for some $\alpha \in I$.

DEFINITION 15.10. A topological space X is **paracompact** if every open cover of X admits a locally finite refinement.

Compact topological spaces are obviously paracompact since a finite subcover can also be viewed as a locally finite refinement. I can now tell you the true reason why we need to assume manifolds are metrizable: *all metrizable spaces are paracompact.* We will not prove quite such a general statement here, but we will make use of the metrizability assumption in the following to prove that manifolds are always paracompact.

LEMMA 15.11. *Every manifold M is σ-compact, i.e. it is the union of countably many compact subsets.*

PROOF. The result is true for every connected locally compact metric space (see e.g. [**Spi99**, Theorem 1.2]), but for our purposes it will be more convenient to drop connectedness and instead assume separability, which holds in any case on all manifolds. Fix a metric d on M that is compatible with its topology. The term "locally compact" refers to the following observation: for every $p \in M$, the closed ball

$$\bar{B}_r(p) := \{ q \in M \mid d(p, q) \leqslant r \}$$

is compact for every $r > 0$ sufficiently small. This holds because whenever r is sufficiently small, $\bar{B}_r(p)$ lies in the domain of a chart that identifies it with a closed and bounded (and therefore compact) subset of Euclidean space. On the other hand, closed and bounded subsets of arbitrary metric spaces are not always compact, so we cannot assume $\bar{B}_r(p)$ is compact for every $r > 0$, but there is a positive (if not infinite) upper bound

$$\kappa(p) := \sup \{ r > 0 \mid \bar{B}_r(p) \text{ is compact} \} \in (0, \infty].$$

If $\kappa(p) = \infty$ at any point p, then M is exhausted by the sequence of compact sets $\bar{B}_k(p)$ for $k = 1, 2, 3, \ldots$ and we are therefore done. Otherwise, observe that by the triangle inequality, every $q \in \bar{B}_{\frac{1}{3}\kappa(p)}(p)$ satisfies

$$\bar{B}_{\frac{1}{3}\kappa(p)}(q) \subset \bar{B}_{\frac{2}{3}\kappa(p)}(p),$$

implying that $\bar{B}_{\frac{1}{3}\kappa(p)}(q)$ is also compact and thus

$$\kappa(q) \geqslant \frac{\kappa(p)}{3} \quad \text{for all} \quad q \in \bar{B}_{\frac{1}{3}\kappa(p)}(p). \tag{15.2}$$

Now for any dense sequence $p_1, p_2, p_3, \ldots \in M$, we claim that

$$M = \bigcup_{k=1}^{\infty} \bar{B}_{\frac{2}{3}\kappa(p_k)}(p_k),$$

where the sets on the right hand side are clearly all compact. Indeed, for any $p \in M$, we can replace $p_1, p_2, p_3, \ldots$ with a subsequence such that $p_k \to p$ as $k \to \infty$, and it follows from (15.2) that $\kappa(p_k) \geqslant \kappa(p)/3$ for all k sufficiently large, so that eventually $p \in \bar{B}_{\frac{2}{3}\kappa(p_k)}$. □

EXERCISE 15.12. Show that if X is a topological space with a locally finite open cover $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ and $K \subset X$ is a compact subset, then K intersects only finitely many of the sets $\mathcal{U}_\alpha$. (It follows from this that if X is σ-compact, then the set I cannot be uncountable, i.e. all locally finite open covers are at most countable. By Lemma 15.11, this applies in particular to all manifolds.

THEOREM 15.13. *Every smooth manifold is paracompact.*

PROOF. Assume $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ is an open cover of M, and using Lemma 15.11, write $M = \bigcup_{j=1}^\infty K_j$ for compact subsets $K_1, K_2, K_3, \ldots$. Choose an open neighborhood $\mathcal{V}_1 \subset M$ of K_1 whose closure is compact, so the set $\bar{\mathcal{V}}_1 \cup K_2$ is also compact. Next, choose $\mathcal{V}_2 \subset M$ to be an open neighborhood of $\bar{\mathcal{V}}_1 \cup K_2$ whose closure is compact, so $\bar{\mathcal{V}}_2 \cup K_3$ is compact. Continuing in this way, one obtains a nested sequence

$$\varnothing =: \mathcal{V}_0 \subset \mathcal{V}_1 \subset \bar{\mathcal{V}}_1 \subset \mathcal{V}_2 \subset \bar{\mathcal{V}}_2 \subset \mathcal{V}_3 \subset \bar{\mathcal{V}}_3 \subset \ldots \subset \bigcup_{j=1}^\infty \mathcal{V}_j = M$$

such that each $\mathcal{V}_j$ is open and each $\bar{\mathcal{V}}_j$ is compact. We will now construct a locally finite refinement of $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ by using the "annular" regions

$$A_j := \bar{\mathcal{V}}_j \backslash \mathcal{V}_{j-1} \subset M, \qquad j = 1, 2, 3, \ldots,$$

which are all compact, and their union is also M. For each $j \in \mathbb{N}$, pick a finite open covering $\{\mathcal{O}^j_\beta \subset M\}_{\beta\in I_j}$ of A_j such that each of the open sets $\mathcal{O}^j_\beta$ is contained in $\mathcal{U}_\alpha$ for some $\alpha \in I$ and is also contained in $\mathcal{V}_{j+1}\backslash\mathcal{V}_{j-2}$. The union of these finite collections for $j = 1, 2, 3, \ldots$ forms an open cover of M that refines $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ and is also locally finite. □

EXERCISE 15.14. Show that without loss of generality, one can assume that all of the open sets in the locally finite refinement given by Theorem 15.13 are diffeomorphic to open balls in Euclidean space.
Remark: This fact is frequently used in proofs that smooth manifolds admit partitions of unity, see for example [**Lee13**, §II.3]. *It is not strictly necessary, however, and we will not use it. The proof given below is conceived to be as close as possible in spirit to proofs of similar results on more general topological spaces, which need not look locally like Euclidean space.*

15.4. Existence of partitions of unity. Now that we know that locally finite refinements can always be found, we need two further ingredients in order to construct partitions of unity. The first is purely topological.

A topological space X is called **normal** if every pair of disjoint closed subsets $A, B \subset X$ have neighborhoods in X that are also disjoint from each other.

EXERCISE 15.15. Show that all metric spaces are normal.

LEMMA 15.16 (the "shrinking lemma"). *Given a locally finite open cover $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ of a normal topological space X, there exists another open cover $\{\mathcal{V}_\alpha\}_{\alpha\in I}$ such that $\bar{\mathcal{V}}_\alpha \subset \mathcal{U}_\alpha$ for every $\alpha \in I$.*

PROOF. We shall give a proof under the extra assumption that the set I is at most countable, which is always true on manifolds due to Exercise 15.12. A proof without this assumption is possible using Zorn's lemma, see e.g. [**nLa**].

Since I is at most countable, we can relable the open cover as $\{\mathcal{U}_i\}_{i=1}^N$ where $N \in \mathbb{N}\cup\{\infty\}$. The sets $A_1 := X\backslash\bigcup_{i=2}^\infty \mathcal{U}_i$ and $X\backslash\mathcal{U}_1$ are closed and disjoint, so we can choose $\mathcal{V}_1 \subset X$ to be any open neighborhood of A_1 that is also disjoint from some neighborhood of $X\backslash\mathcal{U}_1$, implying $\bar{\mathcal{V}}_1 \subset \mathcal{U}_1$. Since $X = \mathcal{V}_1 \cup \bigcup_{i=2}^N \mathcal{U}_i$, we can next take the latter as another open cover on X, and perform the same trick on $\mathcal{U}_2$, producing an open set $\mathcal{V}_2 \subset \bar{\mathcal{V}}_2 \subset \mathcal{U}_2$ such that $X = \mathcal{V}_1 \cup \mathcal{V}_2 \cup \bigcup_{i=3}^N \mathcal{U}_i$. Now repeat this procedure for $i = 3, 4, \ldots, N$, producing a sequence of shrunken open sets $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots \subset X$ such that for each $m \in \mathbb{N}$,

$$X = \bigcup_{i=1}^{m} \mathcal{V}_i \cup \bigcup_{i=m+1}^{N} \mathcal{U}_i. \tag{15.3}$$

If $N < \infty$ then we are done. If $N = \infty$, we now appeal to local finiteness and observe that for every $p \in M$, there exists a largest $m \in \mathbb{N}$ for which $p \in \mathcal{U}_m$, hence (15.3) implies $p \in \bigcup_{i=1}^{m} \mathcal{V}_i$ and thus $X = \bigcup_{i=1}^{\infty} \mathcal{V}_i$. ☐

LEMMA 15.17 (the smooth Urysohn lemma). *Given a smooth manifold M with subsets $A \subset \mathcal{U} \subset M$ such that A is closed and $\mathcal{U}$ is open, there exists a smooth function $f : M \to [0, 1]$ with support in $\mathcal{U}$ such that $f|_A \equiv 1$.*

PROOF, PART 1. For this first of two steps in the proof, we add the assumption that $A \subset M$ is compact. Since the open sets $\mathcal{U}$ and $M \backslash A$ form a finite open cover of M, the compact case of our existence result for partitions of unity (Lemma 11.1) provides a pair of smooth functions $\varphi, \psi : M \to [0, 1]$ that have compact support in $\mathcal{U}$ and $M \backslash A$ respectively such that $\varphi + \psi \equiv 1$ on A. Since $\psi|_A \equiv 0$, the function we were looking for is φ. ☐

Before finishing the proof of Lemma 15.17, it will be convenient to forge ahead and show how these results imply the existence of partitions of unity.

PROOF OF THEOREM 15.3, WITH A CAVEAT. Starting from an arbitrary open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ of M, we can first replace $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ by a locally finite refinement $\{\mathcal{O}_\beta\}_{\beta \in J}$. The latter has the property that for every $\beta \in J$, we can choose some $\alpha(\beta) \in I$ satisfying

$$\mathcal{O}_\beta \subset \mathcal{U}_{\alpha(\beta)}.$$

Next apply the shrinking lemma to find another open cover $\{\mathcal{V}_\beta\}_{\beta \in J}$ such that $\bar{\mathcal{V}}_\beta \subset \mathcal{O}_\beta$ for each $\beta \in J$. By Lemma 15.17, we can choose for each $\beta \in J$ a smooth function $f_\beta : M \to [0, 1]$ with support in $\mathcal{O}_\beta$ such that $f_\beta|_{\bar{\mathcal{V}}_\beta} \equiv 1$. Local finiteness implies that the sum $\sum_{\beta \in J} f_\beta$ is a well-defined smooth function on M, and since every point is contained in at least one of the sets $\mathcal{V}_\beta$, this sum is everywhere positive. Now for each $\alpha \in I$, define $\psi_\alpha : M \to \mathbb{R}$ by

$$\psi_\alpha := \sum_{\{\beta \in J \ | \ \alpha(\beta) = \alpha\}} f_\beta.$$

Local finiteness implies that these are also smooth functions on M and satisfy

$$\sum_{\alpha \in I} \psi_\alpha = \sum_{\beta \in J} f_\beta > 0,$$

and moreover, since each f_β in the sum for $\alpha(\beta) = \alpha$ has support in $\mathcal{O}_\beta \subset \mathcal{U}_\alpha$, ψ_α itself has support in $\mathcal{U}_\alpha$ (see Exercise 15.1). The desired functions φ_α can now be defined by

$$\varphi_\alpha := \frac{\psi_\alpha}{\sum_{\beta \in I} \psi_\beta}.$$

☐

Since we did not yet finish the proof of Lemma 15.17, let's pause now to consider what actually has been proved. Lemma 15.17 was used in the above proof to choose the functions f_β with support in $\mathcal{O}_\beta$ that equal 1 on $\bar{\mathcal{V}}_\beta \subset \mathcal{O}_\beta$. If we add to the hypotheses of Theorem 15.3 that each of the open sets $\mathcal{U}_\alpha \subset M$ has compact closure, then it guarantees that the sets $\bar{\mathcal{V}}_\beta$ are also compact, so that we only need to use the case of Lemma 15.17 that has already been proved. In summary, Theorem 15.3 has now been established under the extra hypothesis that each set $\bar{\mathcal{U}}_\alpha \subset M$ is compact. We can use this observation to complete the proof of Lemma 15.17 and thus establish Theorem 15.3 in full generality.

PROOF OF LEMMA 15.17, PART 2. Choose open coverings $\{\mathcal{U}_\alpha \subset M\}_{\alpha \in I}$ of A and $\{\mathcal{O}_\beta \subset M\}_{\beta \in J}$ of $M \backslash A$ such that all of the sets $\mathcal{U}_\alpha, \mathcal{O}_\beta$ have compact closure and

$$\mathcal{U}_\alpha \subset \mathcal{U} \text{ for all } \alpha \in I, \qquad \mathcal{O}_\beta \subset M \backslash A \text{ for all } \beta \in J.$$

Then $M = \bigcup_{\alpha \in I} \mathcal{U}_\alpha \cup \bigcup_{\beta \in J} \mathcal{O}_\beta$, and we can apply the case of Theorem 15.3 that has been proved already to find a locally finite partition of unity subordinate to this cover: it consists of smooth functions $\{\varphi_\alpha\}_{\alpha \in I}$ and $\{\psi_\beta\}_{\beta \in J}$ such that $\operatorname{supp}(\varphi_\alpha) \subset \mathcal{U}_\alpha$ and $\operatorname{supp}(\psi_\beta) \subset \mathcal{O}_\beta$ for all $(\alpha, \beta) \in I \times J$, while $\sum_{\alpha \in I} \varphi_\alpha + \sum_{\beta \in J} \psi_\beta \equiv 1$. Since every $\mathcal{O}_\beta$ is disjoint from A, it follows that $f := \sum_{\alpha \in I} \varphi_\alpha \equiv 1$ on A, and by Exercise 15.1, $\operatorname{supp}(f) \subset \mathcal{U}$. □

The proof of Theorem 15.3 is now complete.

REMARK 15.18. We made use of separability at one step in this lecture—namely in Lemma 15.11 on σ-compactness—because doing so was more convenient than the alternative, but it was not strictly necessary. As mentioned in the proof of Lemma 15.11, the lemma also holds for arbitrary connected and locally compact metric spaces, so if one works on only one connected component at a time, one obtains a proof of paracompactness for "manifolds" that are assumed metrizable but not necessarily separable. Some authors prefer in fact to define a manifold in a slightly more general way than we have, requiring them to be Hausdorff and paracompact but not necessarily separable or second countable—this shows you how highly the existence of partitions of unity is valued by differential geometers. The only difference this makes in reality is that by the more general definition, manifolds can have uncountably many connected components; in the connected case there is no difference. In any case, I have never seen an example of a non-separable "manifold" that I cared about, not even in infinite dimensions.

REMARK 15.19. On a topological space X, there is generally no well-defined notion of smooth functions, but one can still speak of partitions of unity in which the functions $\varphi_\alpha : X \to [0, 1]$ are only required to be continuous. Such constructions are similarly useful in topology for proving existence results, e.g. the fact that every finite-dimensional *topological* manifold admits a proper topological embedding into $\mathbb{R}^N$ for N sufficiently large (see [**Lee11**, Chapter 4]). To prove that partitions of unity exist on a given space X, one obviously needs to know that X is paracompact, and the other major ingredients are the shrinking lemma (Lemma 15.16) and the continuous variant of Urysohn's lemma (Lemma 15.17), both of which hold whenever X is normal. It turns out that paracompact Hausdorff spaces are automatically normal, thus they admit continuous partitions of unity—in fact for Hausdorff spaces in general, the existence of partitions of unity is *equivalent* to paracompactness.

In nonlinear functional analysis, one sometimes also works with infinite-dimensional smooth manifolds that are locally modelled on Banach spaces. These are not locally compact, so our proof of paracompactness via σ-compactness does not adapt well to the infinite-dimensional setting, but one can nonetheless appeal to the fact that metric spaces are *always* paracompact. The simplest (or at least the shortest) proof of this is due to Mary Ellen Rudin [**Rud69**]. If one considers *arbitrary* metric spaces, then the proof makes slightly mysterious use of the axiom of choice, in the form of the well-ordering theorem: in particular, it uses the fact that for any open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$, the index set I can be endowed with a total order for which every subset has a smallest element. This is less mysterious however in the case of *separable* metric spaces, because every open cover in the separable case admits a finite subcover (exercise!), so one is free without loss of generality to assume the index set is $\mathbb{N}$. As a consequence, infinite-dimensional Banach manifolds are also paracompact, so long as we still agree that anything carrying the name "manifold" should be metrizable and separable. That is the convention that I adopt when I use these objects in my research, and it is not the only possible convention that people might consider reasonable, but it is relatively uncontroversial.

The existence of smooth partitions of unity in the infinite-dimensional setting is nonetheless a subtle question, because smooth compactly-supported "bump" functions do not always exist on Banach spaces—the basic problem here is that on a Banach space E with norm $\|\cdot\|$, the function $E \to \mathbb{R} : x \mapsto \|x\|^p$ is not generally differentiable at $0 \in E$ for any power $p > 0$, even for $p = 2$. As a result, the smooth Urysohn lemma is not true in this context, so smooth partitions of unity do not exist, and many popular constructions from differential geometry are simply not available on infinite-dimensional Banach manifolds. The exception is the case of Hilbert manifolds, which are locally modelled on Hilbert spaces—the inner product on a Hilbert space $\mathcal{H}$ has the convenient property that $\mathcal{H} \to \mathbb{R} : x \mapsto \|x\|^2 := \langle x, x\rangle$ is a smooth function, thus making smooth bump functions and smooth partitions of unity possible.

Exercise 15.20. Given a smooth manifold M, use an open cover and subordinate partition of unity on M to construct a Riemannian metric on the tangent bundle TM. Do *not* assume that Theorem 15.3 holds for TM.
Remark: This exercise ties up a loose end from early in the course: in Corollary 3.12, we defined a smooth structure on the tangent bundle TM of any smooth manifold M, but we never proved that the topology on TM induced by its maximal smooth atlas is metrizable. The existence of a Riemannian metric implies this, and if you follow the instructions in the exercise, its construction does not need to assume that TM is metrizable—it assumes only that M is.

Exercise 15.21. Here is an example of something that satisfies all of the conditions for being a connected smooth 2-manifold except metrizability. It is a variation due to Calabi and Rosenlicht [CR53] on a construction known as the **Prüfer surface**, and can be visualized as a an uncountable collection of planes that have been glued together along their open upper and lower halves, but not along the x-axis, so that the result is a single plane in which the x-axis has been replaced by uncountably many copies of itself. Here is a precise definition: denote the open upper and lower half-planes by $\mathbb{H}_\pm := \{(x,y) \in \mathbb{R}^2 \mid \pm y > 0\}$, and associate to each $a \in \mathbb{R}$ a copy of the full plane $X_a := \mathbb{R}^2$. As a set, the Prüfer surface is

$$\Sigma := \mathbb{H}_+ \cup \mathbb{H}_- \amalg \left(\coprod_{a \in \mathbb{R}} X_a \right) \Big/ \sim$$

where the equivalence relation identifies each point $(x,y) \in X_a$ for $y \neq 0$ with the point $(a + yx, y) \in \mathbb{H}_+ \cup \mathbb{H}_-$. Notice that $\mathbb{H}_\pm$ and X_a for each $a \in \mathbb{R}$ can be regarded naturally as subsets of Σ. Let us denote points $(x,y) \in X_a \subset \Sigma$ by

$$(x,y)_a \in \Sigma,$$

so by definition, $(x,y)_a = (x',y')_b$ whenever $y = y' \neq 0$ and $xy + a = x'y' + b$, but $(x,0)_a$ and $(x',0)_b$ are never equal when $a \neq b$. Prove:

(a) Σ admits a unique smooth structure for which the natural inclusions $\mathbb{H}_\pm \hookrightarrow \Sigma$ and $X_a \hookrightarrow \Sigma$ for each $a \in \mathbb{R}$ are diffeomorphisms onto their images. *Assume for the remaining parts of this exercise that Σ is equipped with the topology uniquely determined by this smooth structure (cf. Prop. 2.12).*

(b) For any two points $p, q \in \Sigma$, there exist neighborhoods $p \in \mathcal{U} \subset \Sigma$ and $q \in \mathcal{V} \subset \Sigma$ such that $\mathcal{U} \cap \mathcal{V} = \varnothing$. (In topological terminology, Σ is Hausdorff.)
Hint: The only case where it is not so obvious is when p and q are both of the form $(x,0)_a$ and $(x',0)_b$. Try drawing pictures of the intersections of neighborhoods of those points with $\mathbb{H}_+ \cup \mathbb{H}_-$.

(c) Σ is connected.
(d) Σ is separable.
Hint: Show that any dense subset of $\mathbb{H}_+ \cup \mathbb{H}_- \subset \Sigma$ *is also dense in* Σ.
(e) Here's where things get weird: the subset $\{(0,0)_a \in \Sigma \mid a \in \mathbb{R}\} \subset \Sigma$ is discrete, i.e. each of its points has a neighborhood that contains none of the others. In particular, all subsets of this set are closed.
(f) Σ is not σ-compact (no pun intended).
Hint: According to part (e), it contains an uncountable discrete subset.

We can now deduce that Σ is not metrizable, as we would otherwise have a contradiction to the proof of Lemma 15.11. Here is an even stranger indication: recall from Exercise 15.15 that all metric spaces are normal.

(g) Suppose we have associated to each $a \in \mathbb{R}$ a "wedge-shaped" region in $\mathbb{H}_+$ of the form

$$W_a := \{(r\cos\theta, r\sin\theta) \in \mathbb{H}_+ \mid r \in (0, r(a)) \text{ and } \theta \in (\pi/2 - \epsilon(a), \pi/2 + \epsilon(a))\}$$

for constants $r(a) > 0$ and $\epsilon(a) > 0$ that are allowed to vary arbitrarily with $a \in \mathbb{R}$. Show that there exists some $a_\infty \in \mathbb{Q}$ and a sequence $a_j \in \mathbb{R}\backslash\mathbb{Q}$ that converges to a_∞ such that $r(a_j)$ and $\epsilon(a_j)$ are both bounded from below.
Big hint: $\mathbb{R} = \mathbb{Q} \cup \bigcup_{N\in\mathbb{N}} A_N$ *where*

$$A_N := \{a \in \mathbb{R}\backslash\mathbb{Q} \mid r(a) \geqslant 1/N \text{ and } \epsilon(a) \geqslant 1/N\}.$$

According to the Baire category theorem, a nonempty complete metric space can never be the countable union of subsets that are nowhere dense, meaning sets whose closures have empty interior. Deduce from this that at least one of the sets A_N *contains an open interval in its closure.*
(h) Deduce that the disjoint subsets

$$Q := \{(0,0)_a \in \Sigma \mid a \in \mathbb{Q}\} \subset \Sigma \qquad \text{and} \qquad I := \{(0,0)_a \in \Sigma \mid a \in \mathbb{R}\backslash\mathbb{Q}\} \subset \Sigma$$

are both closed but do not admit disjoint neighborhoods, i.e. Σ is not normal.
(i) Show that the open cover $\{X_a \subset \Sigma\}_{a\in\mathbb{R}}$ of Σ has no locally finite refinement.
Hint: In any refinement of $\{X_a\}_{a\in\mathbb{R}}$*, points of the form* $(0,0)_a$ *and* $(0,0)_b$ *for* $a \neq b$ *must always belong to different sets in the open cover. Show that for the point* $a_\infty \in \mathbb{R}$ *in part (g), every neighborhood of* $(0,0)_{a_\infty}$ *intersects infinitely many such sets.*

The original Prüfer surface is slightly different from the variation by Calabi and Rosenlicht described above, and can be defined as

$$\Sigma' := \mathbb{H}_+ \amalg \left(\coprod_{a\in\mathbb{R}} X_a \right) \Big/ \sim,$$

where the equivalence relation identifies points $(x,y) \in X_a$ with $(a + yx, y) \in \mathbb{H}_+$ only for $y > 0$. We can visualize Σ' as an uncountable collection of planes that have been glued together along their upper halves, leaving the lower halves separate.

(j) Show that Σ' has all the same properties we proved above for Σ, except that Σ' is not separable.

16. Vector bundles

We have already seen several examples in this course of sets of the form

$$E = \bigcup_{p\in M} E_p,$$

where M is a manifold and E_p is a vector space associated to each point $p \in M$. The obvious example is the tangent bundle TM, but we have also considered the cotangent bundle T^*M, which is the union of the dual spaces to the tangent spaces, and further examples arise in natural ways by thinking of tensor fields $S \in \Gamma(T^k_\ell M)$ as objects that associate to each point $p \in M$ an element S_p of a certain vector space of multilinear maps. For all of these examples, one can regard the vector spaces E_p as "varying smoothly" with respect to p, but this is an intuitive notion that we have not yet made precise except in the special case of TM, on which we defined a smooth structure so that the natural projection $\pi : TM \to M$ sending T_pM to p is a smooth map.

We will now start defining such notions in greater generality.

16.1. Main Definition. We begin with a few more observations about the motivating example of a vector bundle, namely the tangent bundle TM of a smooth n-manifold M. Recall that each chart $(\mathcal{U}, x)$ on M determines a family of vector space isomorphisms

$$d_p x : T_pM \to \mathbb{R}^n, \qquad p \in \mathcal{U}.$$

This information can be repackaged as a bijective map

$$\Phi_x : T\mathcal{U} \to \mathcal{U} \times \mathbb{R}^n$$

whose restriction to each of the individual vector spaces $T_pM \subset T\mathcal{U}$ for $p \in \mathcal{U}$ is $X \mapsto (p, d_px(X)) \in \mathcal{U} \times \mathbb{R}^n$, and the smooth chart $(T\mathcal{U}, Tx)$ for TM can be derived from this by writing

$$Tx(X) = (x(p), d_px(X)) = (x \times \mathbb{1}) \circ \Phi_x(X) \in \mathbb{R}^n \times \mathbb{R}^n \qquad \text{for } X \in T_pM,\ p \in \mathcal{U}.$$

Since $x \times \mathbb{1} : \mathcal{U} \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ is clearly a smooth map, the way that we defined the smooth structure on TM makes Φ_x not just a bijection, but also a diffeomorphism. Now, if $(\mathcal{V}, y)$ is another chart with $\mathcal{U} \cap \mathcal{V} \neq \varnothing$, there is a similar diffeomorphism

$$\Phi_y : T\mathcal{V} \to \mathcal{V} \times \mathbb{R}^n,$$

and both Φ_x and Φ_y restrict to diffeomorphisms $T(\mathcal{U} \cap \mathcal{V}) \to (\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^n$, giving rise to a map

$$\begin{aligned} \Phi_y \circ \Phi_x^{-1} : (\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^n &\to (\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^n \\ (p, v) &\mapsto (p, g(p)v), \end{aligned}$$

where

$$g(p) := d_py \circ (d_px)^{-1} = D(y \circ x^{-1})(x(p)) \in \mathrm{GL}(n, \mathbb{R}) \subset \mathbb{R}^{n \times n}.$$

The smooth compatibility of x and y implies that $g : \mathcal{U} \cap \mathcal{V} \to \mathrm{GL}(n, \mathbb{R})$ is also a smooth function. The existence of maps such as Φ_x and Φ_y is one way of making precise the notion that the tangent spaces T_pM vary smoothly with $p \in M$. We take this as motivation for the definition below.

NOTATION. In everything that follows, we choose a field

$$\mathbb{F} = \text{either } \mathbb{R} \text{ or } \mathbb{C},$$

and assume unless otherwise noted that all vector spaces and linear maps are $\mathbb{F}$-linear. In this way the real and complex cases can be handled simultaneously.

DEFINITION 16.1. Assume M is a smooth n-manifold, E_p is an m-dimensional vector space over $\mathbb{F}$ associated to each point $p \in M$, and define the set

$$E := \bigcup_{p \in M} E_p,$$

where E_p and E_q are regarded as disjoint sets for $p \neq q$.[51] For any subset $\mathcal{U} \subset M$, denote

$$E|_{\mathcal{U}} := \bigcup_{p \in \mathcal{U}} E_p \subset E.$$

A **local trivialization** (*lokale Trivialisierung*) of E is a pair $(\mathcal{U}_\alpha, \Phi_\alpha)$ consisting of an open subset $\mathcal{U}_\alpha \subset M$ and a bijection

$$E|_{\mathcal{U}_\alpha} \xrightarrow{\Phi_\alpha} \mathcal{U}_\alpha \times \mathbb{F}^m$$

such that for each $p \in \mathcal{U}_\alpha$, Φ_α restricts to E_p as a map of the form $v \mapsto (p, \phi_p v)$ for some vector space isomorphism $\phi_p : E_p \to \mathbb{F}^m$.

Any two local trivializations $(\mathcal{U}_\alpha, \Phi_\alpha)$ and $(\mathcal{U}_\beta, \Phi_\beta)$ determine **transition functions** (*Übergangsfunktionen*) $g_{\beta\alpha}, g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathrm{GL}(m, \mathbb{F})$ such that the map $\Phi_\beta \circ \Phi_\alpha^{-1} : (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \mathbb{F}^m \to (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \mathbb{F}^m$ and its inverse take the form

$$\begin{aligned} \Phi_\beta \circ \Phi_\alpha^{-1}(p, v) &= (p, g_{\beta\alpha}(p)v), \\ \Phi_\alpha \circ \Phi_\beta^{-1}(p, v) &= (p, g_{\alpha\beta}(p)v). \end{aligned} \tag{16.1}$$

We say that $(\mathcal{U}_\alpha, \Phi_\alpha)$ and $(\mathcal{U}_\beta, \Phi_\beta)$ are C^k-compatible for $k \in \mathbb{N} \cup \{0, \infty\}$ (or **smoothly compatible** in the case $k = \infty$) if the transition functions $g_{\beta\alpha}$ and $g_{\alpha\beta}$ are of class C^k.

Exercise 16.2. Show that the two transition functions $g_{\alpha\beta}, g_{\beta\alpha} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathrm{GL}(m, \mathbb{F})$ in Definition 16.1 are related to each other by $g_{\beta\alpha}(p) = [g_{\alpha\beta}(p)]^{-1} \in \mathrm{GL}(m, \mathbb{F})$ for all $p \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$, and conclude that $g_{\alpha\beta}$ is of class C^k if and only if $g_{\beta\alpha}$ is.

Remark 16.3. The notion of C^k-compatibility for transition functions is based on the premise that we know what it means to say that a real or complex matrix-valued function on a smooth manifold is of class C^k. This is fine because $\mathbb{R}^{n \times n}$ and $\mathbb{C}^{n \times n}$ can both be regarded as finite-dimensional real vector spaces (every complex vector space is also a real vector space), and the notion of smoothness for functions $f : M \to V$ is well defined whenever M is a smooth manifold and V is a real vector space. The notion of smoothness would be much less clear if we replaced $\mathbb{F}$ with a different field such as $\mathbb{Z}_2$ or $\mathbb{Q}$; there is no theory of differential calculus for functions on open subsets of $\mathbb{R}^n$ with values only in $\mathbb{Z}_2$ or $\mathbb{Q}$. That is one of a few reasons why we will never consider such generalizations in this course.

Definition 16.4. Assume M is a manifold. A **vector bundle of class C^k with rank m over M** (*ein Vektorbündel von der Klasse C^k mit Rang m über M*) is a collection of m-dimensional vector spaces $E = \bigcup_{p \in M} E_p$ as in Definition 16.1, equipped with a maximal collection of C^k-compatible local trivializations $\{(\mathcal{U}_\alpha, \Phi_\alpha)\}_{\alpha \in I}$ such that $M = \bigcup_{\alpha \in I} \mathcal{U}_\alpha$. The vector spaces E_p for $p \in M$ are called the **fibers** (*Fasern*) of the vector bundle E, M is called the **base** (*Basis*) of E, and the set E itself is called the **total space** (*Totalraum*). The surjective map

$$\pi : E \to M$$

sending each fiber $E_p \subset E$ to the point $p \in M$ is sometimes called the **bundle projection**. We will denote the rank of E by

$$\mathrm{rank}_{\mathbb{F}}(E) := m \geqslant 0,$$

or simply $\mathrm{rank}(E)$ whenever the field $\mathbb{F}$ is clear from context.

[51] In set-theoretic terms, this means we are defining E as the disjoint union of all the sets E_p, so we could also have written $E = \coprod_{p \in M} E_p$. We prefer however to avoid the use of the symbol "$\coprod$" here, because we will soon define a topology on E, and it will not be the disjoint union topology.

EXERCISE 16.5. By identifying $\mathbb{C}^m$ with $\mathbb{R}^{2m}$, show that every complex vector bundle E of class C^k can also be regarded as a real vector vector bundle of class C^k with

$$\operatorname{rank}_{\mathbb{R}}(E) = 2\operatorname{rank}_{\mathbb{C}}(E).$$

REMARK 16.6. A vector bundle of rank m is also sometimes called an **m-plane bundle** or an "m-dimensional" vector bundle, and in the case $m = 1$, a **line bundle** (*Geradenbündel*). The latter terminology is quite intuitive when $\mathbb{F} = \mathbb{R}$, but one must keep in mind that in the complex case, the fibers should be visualized as *planes* rather than lines.

NOTATION. We will often refer to the vector bundle in Definition 16.4 simply as E, but doing so ignores quite a lot of important information, such as the base manifold M, fibers E_p, their vector space structures and the local trivializations. It is common in the literature to abbreviate all this data in terms of the projection map and thus refer to $\pi : E \to M$ or (E, π) as a vector bundle, sometimes also omitting the symbol π and writing

$$E \to M.$$

This is an imperfect convention, but we will sometimes also follow it: the projection map has the advantage that it determines the fibers

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

even though it does not determine their vector space structures or the local trivializations.

Observe that if M is a manifold of class C^ℓ for some finite ℓ, then vector bundles of class C^k make sense for every $k \leqslant \ell$, but cannot be defined for $k > \ell$. As usual, we will mostly only consider the case $k = \ell = \infty$, and then refer to E as a **smooth vector bundle**. We also call E a *real* vector bundle if $\mathbb{F} = \mathbb{R}$, and a *complex* vector bundle if $\mathbb{F} = \mathbb{C}$.

REMARK 16.7. The maximal collection of local trivializations $\{(\mathcal{U}_\alpha, \Phi_\alpha)\}_{\alpha \in I}$ in Definition 16.4 plays a similar role to the maximal atlas on a smooth manifold; maximality serves as a bookkeeping device to make sure in this setting that whenever $\{(\mathcal{U}_\alpha, \Phi_\alpha)\}_{\alpha \in I}$ and $\{(\mathcal{V}_\beta, \Psi_\beta)\}_{\beta \in J}$ are two coverings of E by smoothly compatible local trivializations such that every $(\mathcal{U}_\alpha, \Phi_\alpha)$ is smoothly compatible with every $(\mathcal{V}_\beta, \Psi_\beta)$, both can be understood as defining *the same* smooth vector bundle. As with manifolds, one never actually needs to specify a maximal collection of local trivializations, as a maximal collection is uniquely determined by *any* collection $\{(\mathcal{U}_\alpha, \Phi_\alpha)\}$ for which the sets $\mathcal{U}_\alpha$ cover M. When E is a smooth vector bundle, a local trivialization will be called **smooth** whenever it belongs to the associated maximal collection.

REMARK 16.8. Vector bundles of class C^0, also known as *topological* vector bundles, can be defined without assuming the base M is a manifold—the definition makes sense with an arbitrary topological space in place of M, and one can show that E then admits a natural topology such that the bundle projection $\pi : E \to M$ is continuous and the local trivializations are homeomorphisms. (The definition that appears in topology books usually assumes that E is *given* with a topology such that $\pi : E \to M$ is continuous and the fibers $E_p = \pi^{-1}(p)$ are vector spaces; one then calls $\pi : E \to M$ a vector bundle if and only if every $p \in M$ admits a neighborhood $\mathcal{U}$ for which there exists a homeomorphism $\Phi : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times \mathbb{F}^m$ that is a local trivialization.) For many applications, it is also advisable to assume that M is a paracompact Hausdorff space, so that partitions of unity can be used for various constructions, e.g. one can endow the fibers E_p with inner products that depend continuously on p, analogous to a Riemannian metric.

REMARK 16.9. The notion of C^k-compatibility between two local trivializations $(\mathcal{U}_\alpha, \Phi_\alpha)$ and $(\mathcal{U}_\beta, \Phi_\beta)$ could have been defined without mentioning the transition functions $g_{\beta\alpha}, g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathrm{GL}(m, \mathbb{F})$, as it would be equivalent to require that the maps $\Phi_\beta \circ \Phi_\alpha^{-1}$ and $\Phi_\alpha \circ \Phi_\beta^{-1}$ are of class C^k

on $(\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \mathbb{F}^m$. In more general contexts, in particular when one talks in more advanced differential geometry courses about *fiber bundles*, whose fibers are smooth manifolds rather than vector spaces, it becomes necessary to reformulate the notion of smooth compatibility without the transition functions $g_{\alpha\beta}$ and $g_{\beta\alpha}$, as these naturally take values in the diffeomorphism group $\mathrm{Diff}(F)$ of some manifold F, and defining "smoothness" for maps with values in $\mathrm{Diff}(F)$ is something of a technical minefield. We do not have this problem with vector bundles, due to the fact that $\mathrm{GL}(m, \mathbb{F})$ is naturally a smooth finite-dimensional manifold, and (16.1) shows moreover that the transition functions encode all of the essential information in this setting. It will be especially useful to focus on them when we start talking about vector bundles with extra geometric structure, such as bundle metrics or volume forms. In reality, this is also true for most fiber bundles that are of interest, because instead of considering $g_{\alpha\beta}$ and $g_{\beta\alpha}$ with values in $\mathrm{Diff}(F)$, one can often take them to have values in some finite-dimensional smooth Lie group G that acts smoothly on the manifold F.

Here is a generalization of the fact that tangent bundles are smooth manifolds.

Proposition 16.10. *For any smooth vector bundle $\pi : E \to M$ over a smooth manifold M, the total space E naturally has the structure of a smooth manifold of dimension*

$$\dim E = \begin{cases} \dim M + \mathrm{rank}(E) & \text{if } \mathbb{F} = \mathbb{R}, \\ \dim M + 2\,\mathrm{rank}(E) & \text{if } \mathbb{F} = \mathbb{C}, \end{cases}$$

such that the projection map π and the inclusions $E_p \hookrightarrow E$ for $p \in M$ and

$$i : M \hookrightarrow E : p \mapsto 0 \in E_p$$

are all smooth maps.

Proof. The proof is analogous to that of Corollary 3.12, which was the case $E = TM$. The key point is that M can be covered by open sets $\mathcal{U}_\alpha \subset M$ which are domains of charts $x_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^n$ and also appear in local trivializations $\Phi_\alpha : E|_{\mathcal{U}_\alpha} \to \mathcal{U}_\alpha \times \mathbb{F}^m$. The map

$$\phi_\alpha := (x_\alpha \times \mathbb{1}) \circ \Phi_\alpha : E|_{\mathcal{U}_\alpha} \to \mathbb{R}^n \times \mathbb{F}^m \tag{16.2}$$

is then an $(n+m)$-dimensional chart for E on the domain $E|_{\mathcal{U}_\alpha} \subset E$ if $\mathbb{F} = \mathbb{R}$, or if $\mathbb{F} = \mathbb{C}$, an $(n+2m)$-dimensional chart after identifying $\mathbb{C}^m$ with $\mathbb{R}^{2m}$. The smooth compatibility of the charts $(\mathcal{U}_\alpha, x_\alpha)$ and local trivializations $(\mathcal{U}_\alpha, \Phi_\alpha)$ implies (exercise!) that all charts of this form on E are likewise smoothly compatible. The topology defined on E via these charts is metrizable and separable for the same reasons as in the case $E = TM$; in particular, one can use a partition of unity on M to construct a Riemannian metric on the total space E as in Exercise 15.20, proving that E is metrizable. □

Definition 16.11. A **section** (*Schnitt*) of a vector bundle $\pi : E \to M$ is a map $s : M \to E$ such that $\pi \circ s = \mathrm{Id}_M$. In other words, s assigns to each point $p \in M$ a vector in the corresponding fiber $s(p) \in E_p$. We say s is a section **of class** C^k if it is a C^k-map $M \to E$ with respect to the smooth structure on E defined in Proposition 16.10. The vector space of smooth sections is denoted by

$$\Gamma(E) := \left\{ s \in C^\infty(M, E) \;\middle|\; \pi \circ s = \mathrm{Id}_M \right\},$$

with addition and scalar multiplication in $\Gamma(E)$ defined pointwise, e.g. $s + t \in \Gamma(E)$ is defined for $s, t \in \Gamma(E)$ by $(s+t)(p) = s(p) + t(p) \in E_p$.

You might find it unsurprising but not completely obvious that $s + t$ is always a *smooth* section whenever s and t are. To make this obvious, we need to reformulate slightly the meaning of smoothness for a section $s : M \to E$. We observe first that for any local trivialization $\Phi_\alpha : E|_{\mathcal{U}_\alpha} \to \mathcal{U}_\alpha \times \mathbb{F}^m$, every section $s : M \to E$ uniquely determines a vector-valued function

$$s_\alpha : \mathcal{U}_\alpha \to \mathbb{F}^m$$

such that

$$\Phi_\alpha(s(p)) = (p, s_\alpha(p)) \qquad \text{for all } p \in \mathcal{U}_\alpha.$$

We will call this the **local representation** of s with respect to the trivialization $(\mathcal{U}_\alpha, \Phi_\alpha)$. After shrinking $\mathcal{U}_\alpha$ if necessary to a smaller neighborhood of any given point in $\mathcal{U}_\alpha$, we are free to assume that it is also the domain of a chart $x_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^n$, in which case (16.2) defines a corresponding chart $\phi_\alpha : E|_{\mathcal{U}_\alpha} \to \mathbb{R}^n \times \mathbb{F}^m$ for E with the convenient property that its domain contains $s(p)$ for every $p \in \mathcal{U}_\alpha$. Using the charts x_α on M and ϕ_α on E, we obtain a local coordiate representation for the map $s : M \to E$, in the form

$$\phi_\alpha \circ s \circ x_\alpha^{-1} : x(\mathcal{U}_\alpha) \to x(\mathcal{U}_\alpha) \times \mathbb{F}^m : q \mapsto (q, s_\alpha \circ x_\alpha^{-1}(q)).$$

By definition, $s : M \to E$ is a smooth map if and only if this local coordinate representation is smooth for every choice of smooth chart $(\mathcal{U}_\alpha, x_\alpha)$ on M and smooth local trivialization $(\mathcal{U}_\alpha, \Phi_\alpha)$ of E. The latter is clearly true if and only if s_α is a smooth function, so we've proved:

PROPOSITION 16.12. *A section $s : M \to E$ is smooth if and only if its local coordinate representations $s_\alpha : \mathcal{U}_\alpha \to \mathbb{F}^m$ with respect to arbitrary smooth local trivializations $(\mathcal{U}_\alpha, \Phi_\alpha)$ are all smooth.* □

Since $C^\infty(\mathcal{U}_\alpha, \mathbb{F}^m)$ is a vector space for every open set $\mathcal{U}_\alpha$, Proposition 16.12 implies that $\Gamma(E)$ is also a vector space.

EXERCISE 16.13. Show that if $(\mathcal{U}_\alpha, \Phi_\alpha)$ and $(\mathcal{U}_\beta, \Phi_\beta)$ are two local trivializations of E and $s : M \to E$ is a section, then the local representations $s_\alpha : \mathcal{U}_\alpha \to \mathbb{F}^m$ and $s_\beta : \mathcal{U}_\beta \to \mathbb{F}^m$ are related to each other on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ in terms of the transition function $g_{\beta\alpha} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathrm{GL}(m, \mathbb{F})$ by

$$s_\beta(p) = g_{\beta\alpha}(p) s_\alpha(p) \qquad \text{for } p \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta.$$

Remark: Since the transition functions on a smooth vector bundle are all smooth, this exercise implies that the condition in Proposition 16.12 does not need to be checked for all possible smooth local trivializations—it suffices to check it for a family of trivializations that cover M.

DEFINITION 16.14. Assume $E \to M$ and $F \to M$ are two smooth vector bundles over the same manifold M. A smooth map $\Psi : E \to F$ is called a **smooth linear bundle map** if for every $p \in M$, the restriction $\Psi|_{E_p}$ is a linear map

$$\Psi_p : E_p \to F_p.$$

We call Ψ **fiberwise injective** / **surjective** if Ψ_p is injective / surjective for every $p \in M$, and Ψ is a **bundle isomorphism** if Ψ_p is a vector space isomorphism for every $p \in M$. The bundles E and F are called **isomorphic** if and only if there exists a bundle isomorphism $E \to F$.

REMARK 16.15. Definition 16.14 presumes that E and F are both bundles over the same field $\mathbb{F}$. If one is a real vector bundle and the other is complex, then one can always regard the complex bundle as a real bundle with twice the rank (see Exercise 16.5) and thus interpret $\Psi : E \to F$ as a smooth *real*-linear bundle map.

EXERCISE 16.16. Suppose $E, F \to M$ are smooth vector bundles and $\Psi : E \to F$ is a map whose restriction to E_p for each p is a linear map $\Psi_p : E_p \to F_p$.

(a) Show that for every pair of smooth local trivializations $\Phi_\alpha : E|_{\mathcal{U}_\alpha} \to \mathcal{U}_\alpha \times \mathbb{F}^m$ and $\Psi_\beta : E|_{\mathcal{U}_\beta} \to \mathcal{U}_\beta \times \mathbb{F}^k$, there exists a unique function

$$\Psi_{\beta\alpha} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathbb{F}^{k\times m}$$

such that

$$\Phi_\beta \circ \Psi \circ \Phi_\alpha^{-1} : (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \mathbb{F}^m \to (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \mathbb{F}^k : (p, v) \mapsto (p, \Psi_{\beta\alpha}(p)v).$$

(b) Show that Ψ is a smooth linear bundle map if and only if for all choices of the two smooth local trivializations in part (a), the function $\Psi_{\beta\alpha}$ is smooth.

Definition 16.17. Given a smooth vector bundle $E \to M$, a **smooth subbundle** (*Unterbündel*) of E is a vector bundle $F \to M$ such that for each $p \in M$, F_p is a linear subspace of E_p, and the inclusion $F \hookrightarrow E$ is a smooth linear bundle map.

16.2. Some basic examples. We now relate the definitions above to various examples that have already appeared in this course. For several of them, there is still some work to be done in showing that they naturally admit coverings by families of smoothly compatible local trivializations, and this work will be postponed until the next lecture.

Example 16.18 (**tangent bundle**). If M is an n-manifold, its tangent bundle TM is a smooth real vector bundle of rank n, where each smooth chart $(\mathcal{U}, x)$ determines a local trivialization $\Phi : TM|_{\mathcal{U}} = T\mathcal{U} \to \mathcal{U} \times \mathbb{R}^n$ by $\Phi(X) = (p, d_p x(X))$ for $X \in T_pM$. A smooth section of TM is nothing other than a smooth vector field on M,

$$\Gamma(TM) = \mathfrak{X}(M).$$

Example 16.19 (**cotangent bundle**). The cotangent bundle T^*M of a smooth n-manifold M has fibers $T_p^*M = \operatorname{Hom}(T_pM, \mathbb{R})$ for $p \in M$. We will construct smoothly compatible local trivializations for T^*M in the next lecture—it is a special case of the fact that every smooth vector bundle has a *dual bundle* which is also a smooth vector bundle in a natural way. The smooth sections of T^*M will then be the smooth 1-forms on M,

$$\Gamma(T^*M) = \Omega^1(M).$$

Example 16.20 (**tensor and exterior bundles**). For each $k, \ell \geqslant 0$, there is a natural smooth real vector bundle $T_\ell^k M \to M$ of rank $n^{k+\ell}$ whose fiber at a point p is the vector space $(T_pM)_\ell^k$ of $(k+\ell)$-fold multilinear maps $T_p^*M \times \ldots \times T_p^*M \times T_pM \times \ldots \times T_pM \to \mathbb{R}$. The smooth local trivializations on $T_\ell^k M$ will also arise from more general constructions to be discussed in the next lecture. Consistently with our previous notation, the space of smooth sections $\Gamma(T_\ell^k M)$ will then be precisely the space of smooth tensor fields of type (k, ℓ).

For each $k \geqslant 0$, there is an important subbundle

$$\Lambda^k T^*M \subset T_k^0 M$$

of rank $\binom{n}{k}$ whose fiber over $p \in M$ is the vector space of alternating k-forms $\Lambda^k T_p^*M \subset (T_pM)_k^0$. The sections of $\Lambda^k T^*M$ will of course be the smooth differential k-forms,

$$\Gamma(\Lambda^k T^*M) = \Omega^k(M).$$

Note that by definition,

$$T_1^0 M = T^*M = \Lambda^1 T^*M,$$

and since $(T_pM)_0^0$ is defined simply as $\mathbb{R}$ for every p, $T_0^0 M = \Lambda^0 T^*M$ is simply the *trivial* line bundle $M \times \mathbb{R}$ (cf. Example 16.21 below).

EXAMPLE 16.21 (**trivial bundle**). For any manifold M, the **trivial m-plane bundle** over M is the product $E = M \times \mathbb{F}^m$, with fibers

$$E_p := \{p\} \times \mathbb{F}^m,$$

understood in the obvious way as m-dimensional vector spaces. This is a smooth vector bundle because (M, Id) is a local trivialization that covers the entirety of M, so the associated maximal collection of local trivializations consists of all that are smoothly compatible with this one. Smooth sections $s : M \to M \times \mathbb{F}^m$ are smooth maps of the form $p \mapsto (p, f(p))$ and are thus equivalent to smooth functions $f : M \to \mathbb{F}^m$.

DEFINITION 16.22. A vector bundle $\pi : E \to M$ of rank m is (globally) **trivial**[52] if it admits a bundle isomorphism to the trivial m-plane bundle over M.

A local trivialization $\Phi : E|_{\mathcal{U}} \to \mathcal{U} \times \mathbb{F}^m$ of a vector bundle E can be understood as a bundle isomorphism between the restriction $E|_{\mathcal{U}} \to \mathcal{U}$ and the trivial m-plane bundle over $\mathcal{U}$. By definition, every vector bundle is therefore *locally* trivial, meaning that its restriction to any sufficiently small open subset must be trivial. The next example shows that globally, this need not be true.

EXAMPLE 16.23 (**a nontrivial real line bundle**). Identify S^1 with the unit circle in $\mathbb{C}$, and define $\ell \subset S^1 \times \mathbb{R}^2$ as the union of the sets $\{e^{i\theta}\} \times \ell_{e^{i\theta}} \subset S^1 \times \mathbb{R}^2$ for all $\theta \in \mathbb{R}$, where $\ell_{e^{i\theta}} \subset \mathbb{R}^2$ is the 1-dimensional subspace

$$\ell_{e^{i\theta}} = \mathbb{R}\begin{pmatrix}\cos(\theta/2)\\ \sin(\theta/2)\end{pmatrix} \subset \mathbb{R}^2.$$

Exercise 16.24 below shows that ℓ can be regarded as a smooth line bundle over S^1 with fibers $\ell_{e^{i\theta}}$ for $e^{i\theta} \in S^1$. Observe that if we consider the subset

$$\{(e^{i\theta}, v) \in \ell \mid \theta \in \mathbb{R},\ |v| \leqslant 1\}$$

consisting only of vectors of length at most 1, we obtain a Möbius strip. Local trivializations of $\ell \to S^1$ can be constructed as follows: for any $\theta_0 \in \mathbb{R}$, set $p := e^{i\theta_0} \in S^1$, and define

$$\Phi : \ell|_{S^1\setminus\{p\}} \to (S^1\setminus\{p\}) \times \mathbb{R} : \left(e^{i\theta}, c\begin{pmatrix}\cos(\theta/2)\\ \sin(\theta/2)\end{pmatrix}\right) \mapsto (e^{i\theta}, c), \tag{16.3}$$

with θ assumed to vary in the interval $(\theta_0, \theta_0 + 2\pi)$.

EXERCISE 16.24. For the line bundle $\ell \to S^1$ in Example 16.23, prove:

(a) Any two local trivializations defined as in (16.3) with different choices of $\theta_0 \in \mathbb{R}$ are smoothly compatible.
(b) ℓ is a smooth subbundle of the trivial bundle $S^1 \times \mathbb{R}^2$.
(c) There exists no continuous section of ℓ that is nowhere zero.
(d) ℓ is not globally trivial.

[52] If we were being more pedantic, we would say **globally trivializable** in Definition 16.22 instead of "trivial", and reserve the latter for any vector bundle that is literally presented as a product $M \times \mathbb{F}^m$ with the identity map as a smooth trivialization, rather than just being isomorphic to one. But the looser use of the word "trivial" to mean "isomorphic to a trivial bundle" is widespread, so you should get used to it.

Bibliography

[AF03] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, 2nd ed., Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[CR53] E. Calabi and M. Rosenlicht, *Complex analytic manifolds without countable base*, Proc. Amer. Math. Soc. **4** (1953), 335–340.

[DK90] S. K. Donaldson and P. B. Kronheimer, *The geometry of four-manifolds*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1990. Oxford Science Publications.

[Gro85] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), no. 2, 307–347.

[Hir94] M. W. Hirsch, *Differential topology*, Springer-Verlag, New York, 1994.

[Lee11] J. M. Lee, *Introduction to topological manifolds*, 2nd ed., Graduate Texts in Mathematics, vol. 202, Springer, New York, 2011.

[Lee13] ———, *Introduction to smooth manifolds*, 2nd ed., Graduate Texts in Mathematics, vol. 218, Springer, New York, 2013.

[nLa] nLab, *Shrinking lemma*. exposition based on a blog post by Matt Rosenzweig, available at https://ncatlab.org/nlab/show/shrinking+lemma.

[Rud69] M. E. Rudin, *A new proof that metric spaces are paracompact*, Proc. Amer. Math. Soc. **20** (1969), 603.

[Sal16] D. A. Salamon, *Measure and integration*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2016. MR3469972

[Spi99] M. Spivak, *A comprehensive introduction to differential geometry*, 3rd ed., Vol. 1, Publish or Perish Inc., Houston, TX, 1999.

[Wen23] C. Wendl, *Topology I and II* (2023). Notes from the course at HU Berlin, available at https://www.mathematik.hu-berlin.de/~wendl/Winter2023/Topologie2/lecturenotes.pdf.

[Wen19] ———, *Integration auf Untermannigfaltigkeiten* (2019). Skript zur Vorlesung Analysis III an der HU Berlin, verfügbar unter https://www.mathematik.hu-berlin.de/~wendl/Winter2019/Analysis3/Skript_DifferentialFormen.pdf