These course-notes are a minor revision (from 2003) of a first draft that was prepared for a course in spring 2000 at ASU. They are extremely close to (but nowhere as accurate) as Spivak’s books – and they can be justified only by Spivak being out of print when the 2000 spring semester began. Since then Spivak’s volumes have been republished (now typeset, no longer type-written), and every user of these notes is expected to eventually buy the originals by Spivak – he deserves his royalties!

These notes are based on class-notes taken from a class taught by Al Lundell at the University of Colorado in 1983/84, and on two classes taught in 1989 and 1991 at Arizona State University. Originally, these notes may have been quite independent, but upon efforts to make them more comprehensive, more precise, they have again converged very much to Spivak’s treatment.

However, in some places the presentation departs from Spivak and material follows more closely e.g. Sternberg (e.g. notation and terminology involving tensors), Boothby (e.g. Riemannian basics), Marsden and others. The main practical value of these notes is that they use the same notation (even if it is just $u^i$ in place of $x^i$) that the instructor has become too accustomed to, and will use in class...

The current version does not include any diagrams – which are essential for readability. Moreover, sections such as the reviews of basic topology and differentiability, which really belong into an appendix, are included in the order that the class actually covered them.

The affiliated explorations that use computer algebra system have not yet been integrated into these notes.
1 Curves in the plane and 3-space

This first section addresses mostly prerequisite material and is not completely self-contained. It provides some basic definitions and discusses some fundamental theorems. Central objectives are to raise some questions that will have to be addressed when working in more general settings, and to set the stage for the questions about geometric properties.

1.1 Basic definition of a curve

In many settings it may be appropriate to think of a curve as a set of points in the plane or in 3-space. However, in differential geometry and other advanced settings, it is generally more convenient to work with a different notion – basically calling what previously was named a parameterization the “curve”.

Definition 1.1 A curve is a continuous function defined on an interval $I \subseteq \mathbb{R}$, taking values in a (topological) space $M$ (in this section $M$ is assumed to be $\mathbb{R}^n$). (The interval in this definition may be open or closed, finite, semi-infinite or the entire real line. At this time we only assume enough structure on the space $M$ so that we can talk about continuity.)

The key difference is that with this definition a curve is a function. Consequently it has a richer structure than just a set of points – a structure that facilitates technical analysis. Moreover, this definition easily carries over to much more general settings – e.g. we may think of a vibrating membrane as a curve in an appropriate space $M$ of functions of two variables. What matters is that the space has enough structure (at least a topology) so that we may talk about continuity. (Later we will require additional structures on the space $M$ so that we can differentiate curves.)

Several properties of curves deserve their own names. A curve $\gamma: I \mapsto M$ is called closed if $I = [a, b]$ is a (finite) closed interval and $\gamma(a) = \gamma(b)$. If the restriction of a closed curve $\gamma$ to $[a, b]$ is one-to-one, then $\gamma$ is called a simple closed curve.

Example 1.1 The circle $S^1 = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 = 1\}$ is the image of the simple closed curve $\gamma: [0, 2\pi] \mapsto \mathbb{R}^2$ defined by $\gamma(t) = (\cos t, \sin t)$. (Note that the circle $S^1$ is not a curve!)

Definition 1.2 If $\gamma: I \mapsto M$ is a curve defined on a finite interval $I$ and $\phi: J \mapsto I$ is a continuous function that maps a finite interval $J$ onto $I$ (mapping endpoints to endpoints), then the curve $\sigma = \gamma \circ \phi: J \mapsto M$ is called a reparameterization of $\gamma$.

The notion of reparameterization may be extended to infinite intervals provided one suitably modifies the notion of endpoint to endpoint, e.g. requiring the existence of the limits $\lim_{t \to \pm\infty}$ and that these equal $-\infty$ and $\infty$. Among one-to-one reparameterizations one distinguishes orientation-preserving and orientation-reversing reparameterizations according to whether the map $\phi$ maps the left endpoint of $J$ to the left or right endpoint of $I$. 

Picture: $I$, $J$, $\gamma(I)$ and arrows for $\gamma, \phi, \sigma$
1.2 Differentiable curves and arc-length

An intuitive notion of the length of a curve in $\mathbb{R}^n$ may be built on successive approximations by polygonal approximations. More specifically, suppose that $\gamma: [a, b] \to \mathbb{R}^n$. Let $\| \cdot \|$ denote the Euclidean norm $\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \ldots + x_n^2}$ in $\mathbb{R}^n$. Define the length $L(\gamma)$ of the curve as the supremum (possibly infinite)

$$
L(\gamma) = \sup_{\mathcal{P}} \sum_{j=0}^{n(\mathcal{P})} \| \gamma(t_{i+1}) - \gamma(t_i) \|
$$

(1)

where $\mathcal{P}$ ranges over all partitions $\mathcal{P} = \{ t_i: 0 \leq i \leq n(\mathcal{P}) \}$ such that $a = t_0 < t_1 < \ldots < t_{n(\mathcal{P})} = b$.

It is very important to note that this definition of length cannot directly generalize to spaces for which one does not have an a-priori notion of distance – i.e. where $\| \cdot \|$ has no meaning (yet).

The key idea is that for differentiable curves there is a natural alternative – the length is the integral of the speed, and this notion will generalize, even give rise to the concept of Riemannian manifolds. Loosely speaking, the main idea is to rewrite

$$
\sum_{j=0}^{n} \| \gamma(t_{i+1}) - \gamma(t_i) \| = \sum_{j=0}^{n} \frac{\| \gamma(t_{i+1}) - \gamma(t_i) \|}{(t_{i+1} - t_i)} \cdot (t_{i+1} - t_i) \to \int_{a}^{b} \| \gamma'(t) \| \, dt =: L(\gamma) \quad (2)
$$

By fairly straightforward (advanced) calculus arguments one may make this idea rigorous, i.e. show that for any continuously differentiable curve defined on a finite closed interval there exists a unique limit which defines the length of the curve.

The key to most of our later work will be to develop a natural notion of a tangent vector to a curve (taking values in an abstract manifold) that does not require any prior notion of a difference $\gamma(t_{i+1}) - \gamma(t_i)$ of two points in that space. Once we have such a generalized notion of a tangent vector, much of the following fundamental notions, structures, calculations and arguments will carry over to the general case of abstract manifolds.

**Definition 1.3** A curve $\gamma: (a, b) \to \mathbb{R}^n$ is called differentiable if for every $t \in (a, b)$ the limit $\lim_{h \to 0} \frac{1}{h} (\gamma(t + h) - \gamma(t))$ exists. If the limit exists, it is denoted $\gamma'(t)$ and called the velocity at $\gamma(t)$ (or at $t$).

The second derivative $\gamma''(t)$ is defined analogously, and is usually called the acceleration at $\gamma(t)$ or at $t$. The magnitude $\| \gamma'(t) \|$ of the velocity is called the speed.

In the case of plane and space curves one routinely identifies the point $\gamma(t) \in \mathbb{R}^n$ with the arrow (vector) from the origin to this point. On the other hand the velocity and acceleration are commonly visualized as arrows (vectors) rooted at the point $\gamma(t)$. There appears to be a certain arbitrariness about this representation – but it seems to make sense after a little thought. The upcoming construction of the tangent bundle will illuminate the situation and provide clarifying distinctions. A helpful preparation at this time is to think about possible alternative representations, and to find good arguments why the usual placements of the arrows are a good choice without any compelling alternative.

**Exercise 1.1** Show that if the velocity $\gamma'$ is constant then the (image of the) curve $\gamma$ is a straight line, but the converse is not true.

**Exercise 1.2** Verify that the curve $\gamma: \mathbb{R} \to \mathbb{R}^2$ defined by $\gamma(t) = (\cos(t), \sin(t))$ has constant speed, but nonzero acceleration.

Exercise 1.3 Prove that the acceleration \( \gamma'' \) is orthogonal to the velocity \( \gamma' \) (for all \( t \)), i.e. \( <\gamma'(t), \gamma''(t)> = 0 \) if and only if the speed is constant.

(Hint: Differentiate \( 0 = ||\gamma'(t)||^2 = <\gamma'(t), \gamma'(t)> \). Read the identities both directions.)

Exercise 1.4 Verify that the curve \( \gamma: \mathbb{R} \mapsto \mathbb{R}^2 \) defined by \( \gamma(t) = (t^2, 0) \) if \( t \geq 0 \), and \( \gamma(t) = (0, t^2) \) if \( t < 0 \) is continuously differentiable, yet its image in the plane has a corner.

**Blanket assumption:**

For most of the following we shall assume that all curves under consideration are at least twice continuously differentiable and that \( ||\gamma'(t)|| \neq 0 \) for all \( t \).

In many cases we will for convenience even assume that the curve is smooth, i.e. that it has continuous derivatives of all orders. Note that the assumption that the speed is never zero eliminates such nuisances as the corners exhibited by the continuously differentiable curve of exercise 1.4. Moreover, it also prohibits such nuisances as exhibited by the curve \( \gamma(t) = (\cos(t - t^3), \sin(t - t^3)) \) which “go back and fourth” along the image of the curve. If there is a need to allow for such behaviours it is usually easy to either consider the curve in pieces, or relax the requirements in specific cases and then adapt the desired theorems as needed.

Obviously many different curves may have the same image – and there seems an arbitrariness about picking a specific parameterization. For theoretical purposes it is often convenient to work with a canonical reparameterization of a differentiable curve \( \gamma \). One such natural choice is such that the parameter \( t \) (often considered as time) agrees with the distance traveled along the curve.

**Definition 1.4** The arc-length of a continuously differentiable curve \( \gamma: [a, b] \mapsto \mathbb{R}^n \) is defined as the function \( s: [a, b] \mapsto \mathbb{R} \),

\[
s(t) = \int_a^t <\gamma' (\tau), \gamma' (\tau)>^{1/2} d\tau \quad (3)
\]

The curve \( \gamma \) is called parameterized by arc-length if \( s(t) \equiv t \), or, equivalently, \( <\gamma', \gamma'> \equiv 1 \).

This definition relies on the standard inner product \( <\cdot, \cdot> \) which in \( \mathbb{R}^n \) is almost synonymous with the notion of (Euclidean) distance. The key idea underlying Riemannian geometry – see chapter 4 – is that once one has a suitable generalization of this inner product, then most of the notions and properties naturally carry over to abstract settings.

To explicitly reparameterize a given curve \( \gamma: [a, b] \mapsto \mathbb{R}^n \) by arc-length generally requires one not only to evaluate the integral in equation (3) in closed form, but in addition, to solve this equation for \( t \) in terms of \( s \) – generally a hopeless task if looking for closed form expressions in terms of the traditional elementary functions. (Thus one usually has to choose between either formal expressions (for theoretical purposes), and numerical techniques (in practical calculations). We will see that curves that are parameterized by arc-length allow for particular elegant descriptions of their geometry.

**Exercise 1.5** Reparameterize the curve \( \gamma: [0, 1] \mapsto \mathbb{R}^2 \) defined by \( \gamma(t) = (\sqrt{1 - t^2}, t) \) by arc-length by explicitly integrating equation (3) and solving for \( t \) in terms of \( s \).

**Exercise 1.6** Calculate and sketch the graphs of the arc-lengths of the curves \( \gamma_1(t) = (t, t) \), \( \gamma_2(t) = (\cos 2nt, \sin 2nt) \), \( \gamma_3(t) = (t, t^2) \), and \( \gamma_4(t) = (\cos 2\pi t, \sin 2\pi t, ct) \), all defined for \( t \geq 0 \). If feasible, reparameterize each curve by arc-length.
Exercise 1.7 Verify by direct calculation that arc-length of plane curves is invariant under orthogonal linear transformations: More specifically, let $\gamma: [a, b] \mapsto \mathbb{R}^2$ and $\sigma: [a, b] \mapsto \mathbb{R}^2$ be two differentiable curves that are related by $\sigma = A \cdot \gamma$ where $A$ is a $2 \times 2$ rotation matrix with $a_{11} = \pm a_{22} = \cos \theta$ and $\mp a_{12} = a_{21} = \sin \theta$ for some value of $\theta \in \mathbb{R}$. Calculate and compare the arc-length functions associated to $\sigma$ and $\gamma$.

Repeat for reflections with $a_{11} = -a_{22} = \cos \theta$ and $a_{12} = a_{21} = \sin \theta$.

Explain in geometric terms - e.g. using the earlier definition in terms of polygonal approximations – why this is expected. Try to make this into a rigorous argument that applies to $\mathbb{R}^n$ for any dimension $n \geq 1$.

1.3 Curvature of plane and space curves

Curvature is the central concept of differential geometry. In the case of graphs of functions $y = f(x)$ all calculus students learn that the second derivative is somehow related to how much the graph curves – but it is important to fully understand that, and why, the second derivatives does not represent curvature.

Exercise 1.8 Consider the graph of $f_0(x) = x^2$ for $0 \leq x \leq b$. For small angles $\theta$ and small values of $b > 0$ the image of the graph under a rotation by an angle $\theta$ about the origin is again the graph of a function $f_\theta(x)$. Find an explicit formula for $f_\theta$ and show that its second derivative is not constant equal to $f_\theta'' = 2$.

Suggestion: Use $x$ and $y$ to denote points on the original curve $y = x^2$ and let $\xi$ and $\eta$ denote points on the rotated curve. Express $x$ and $y$ in terms of $\xi$ and $\eta$ (compare exercise 1.7), substitute into $y = x^2$ and solve for $\eta$ in terms of $\xi$. Finally calculate $\frac{d^2 \eta}{d\xi^2}$.

See also the associated MAPLE worksheet.

There are two aspects of the second derivative that do not make it suitable for immediate use to denote a notion of curvature: First the derivative in the exercise is taken with respect to the first coordinate $x$, as opposed to the intrinsic arc-length parameter. Secondly, the slopes are not the same as the direction of the curve – the tan in $y' = \tan \alpha$ distorts the description.

Thus we first consider smooth curves $\gamma: I \mapsto \mathbb{R}^2$ and $\gamma: I \mapsto \mathbb{R}^3$ that are parameterized by arc-length. This implies that $\| \gamma'(s) \| = 1$, i.e. the velocity is a unit tangent vector to the curve at any time $s \in I$, suggesting the notation $T(s)$ for $\gamma'(s)$. In the case of plane curves it is convenient to complete $\{T(s)\}$ to a positively oriented orthonormal basis $\{T(s), N(s)\}$.

To describe the rate of change of this direction differentiate again. We define the curvature to be the (signed) magnitude $\kappa = \| \gamma'' \|$ of this derivative. More specifically, in the case of plane curves it is convenient to define $\kappa(s) = \langle T'(s), N(s) \rangle$ (allowing both positive and negative values). In the case of space curves, which will be considered from here on, the natural way of choosing a direction for a normal $N$ is to require that the curvature is nonnegative and use $\gamma''(s) = \kappa(s) N(s)$ as the defining equation for both $\kappa(s)$ and $N(s)$ – of course, in the case that $\| \gamma'' \| = 0$ this only defines $\kappa(s)$, but determines no direction $N(s)$. Recall from exercise 1.3 that $\langle \gamma', \gamma' \rangle \equiv 1$ implies that $\gamma'' \perp \gamma'$, and hence $N(s) \perp T(s)$ for each $s \in I$ such that $\kappa(s) \neq 0$.

Before differentiating further, define for each $s \in I$ where $\kappa(s) \neq 0$ a third unit vector $B(s) = T(s) \times N(s)$ (using the standard cross-product in 3-space). (Observe the analogy to predetermining $N$ in the planar case as the last vector to complete an orthonormal frame – and allowing the coefficient to have both positive and negative values. This also suggests an obvious generalization to higher dimensions $n \geq 3$.) The triple $(T(s), N(s), B(s))$ is called the Frenet frame along the curve $\gamma$ – it is only defined at points where $\kappa(s) \neq 0$.
To continue the investigations differentiate \( N \) (with respect to \( s \)). By an argument analogous to the one employed earlier, \( \frac{d}{ds} N(s) \) is orthogonal to \( N(s) \), and hence may be written as a linear combination \( \frac{d}{ds} N(s) = a_{21}(s) T(s) + \tau(s) B(s) \). The coefficient \( \tau \) is called the torsion of the curve. Intuitively, the torsion quantifies the rate at which the curve twists out of a plane – compare the exercises.

To identify the parameter \( a_{21}(s) \), differentiate the identity \( 0 \equiv <T(s), N(s)> \) and find that \( 0 \equiv \kappa(s) <N(s), N(s)> + a_{21}(s) <T(s), T(s)> + \tau(s) <T(s), B(s)> \). Since the third term vanishes, it is clear that \( a_{21}(s) = -\kappa(s) \). To complete the analysis differentiate the identities \( 0 \equiv <T(s), B(s)> \), \( 0 \equiv <N(s), B(s)> \), and \( 0 \equiv <N(s), N(s)> \) to obtain \( \frac{d}{ds} B(s) = -\tau(s) N(s) \).

Taken together, these equations form the famous Frenet-Serret formulas:

\[
\begin{align*}
T' & = \kappa N \\
N' & = -\kappa T + \tau B \\
B' & = -\tau N
\end{align*}
\]

To emphasize the characteristic structure of this set of equations, we rewrite it by forming a matrices \( R = (T \mid N \mid B) \) and \( R' = (T' \mid N' \mid B') \) whose columns are the representations of these vectors with respect to the standard coordinates in \( \mathbb{R}^3 \). Note that \( R \) is an orthogonal matrix, i.e. \( R^T R = R R^T = I_{3\times3} \). With this notation, the Frenet-Serret formulas become

\[
R' = RA \quad \text{where} \quad A = \begin{pmatrix}
0 & -\kappa & 0 \\
\kappa & 0 & -\tau \\
0 & \tau & 0
\end{pmatrix}
\]

Preview: We will later consider \( R \) as a point (!) in the three dimensional manifold (actually a Lie group) \( SO(3) \) of \( 3 \times 3 \) orthogonal matrices. The product \( RA \), which again is a \( 3 \times 3 \) orthogonal matrix (though generally not orthogonal) will be considered a tangent vector at \( R \), and it will be clear from general considerations (in some sense generalizing the arguments made above about the derivatives of the inner products between the columns of the matrix \( R \)) why \( A \) has to be skew symmetric, i.e. \( A^T = -A \).

When given an initial reference frame \( R(0) \) consisting of three orthogonal unit vectors \( T(0), N(0) \) and \( B(0) \) together with two sufficiently regular functions \( \kappa(s) \) and \( \tau(s) \) this system of differential equations uniquely determines \( R(t) \) for all times! (We will consider \( R \) as a curve in \( SO(3) \).)

Continuing further, if in addition an initial point \( \gamma(0) \in \mathbb{R}^3 \) has been specified, then the Frenet Serret formulas together with the differential equation \( \gamma' = T(s) \) uniquely determine a curve in \( \mathbb{R}^3 \). It is straightforward to verify that the curvature and torsion of this curve agree with the data provided to the differential equation. A most important corollary of this study is that the curvature and torsion completely determine a smooth curve up to translation (as determined by \( \gamma(0) \)) and rotation (as determined by \( R(0) \)).

**Exercise 1.9** Explain how this gives as a corollary that curvature and torsion are invariant under translation and rotation (compare also exercise 1.7).

**Exercise 1.10** Explore how complicated a brute force calculation is (similar to exercise 1.7) that directly shows that curvature and torsion are invariant under rotations and translation. (It may be appropriate to use a computer algebra system for part of this work.)

**Exercise 1.11** Show that if the curvature \( \kappa(s) \equiv 0 \) (for \( s \in [a, b] \)) of a plane curve vanishes identically, then the curve is a straight line. Is the same true for a curve in 3-space? Explain!
Exercise 1.12 Show that if the curvature $\kappa(s) \equiv c \neq 0$ (for $s \in [a, b]$) of a plane curve is constant, then the curve is a circle with radius $1/c$.

Is the same true for a curve in 3-space? Explain! (Remark: Feel free to consult the literature for elegant arguments – a direct brute-force approach quickly can get very messy!)

Exercise 1.13 Show that if the torsion $\tau(s) \equiv 0$ (for $s \in [a, b]$) of a space curve vanishes identically, then the curve lies in a plane.

The Frenet-Serret formulas provide a most beautiful and comprehensive geometric description of the curves in 3-space. They appear to intrinsically rely on working with parameterizations by arc-length, yet for most curves explicit closed-form formulas for parameterizations by arc-length are beyond reach. However, note that all these formulas only involve derivatives of the curve $\gamma$.

Consequently, there is no need to ever explicitly calculate the arc-length. All that is needed is the integrand of the formula (3) – the chain-rule does the rest as shown below!

Consider a smooth curve $\sigma: J = [a, b] \mapsto \mathbb{R}^n$. Define $\phi(t) = \int_a^t <\sigma'(\tau), \sigma'(\tau)> \, d\tau$. As usual, assume that $\|\gamma'(t)\| > 0$ for all $t \in I$. Then the curve $\gamma: J = [0, L(\sigma)] \mapsto \mathbb{R}^n$ defined by $\gamma = \sigma \circ \phi^{-1}$ is the reparameterization of $\sigma$ by arc-length.

The chain rule relates the velocities $\sigma'(t) = \gamma'(\phi(t)) \phi'(t) = T(\phi(t)) \cdot \|\sigma'(t)\|$, $\sigma''(t) = \gamma''(\phi(t)) \phi''(t) + \gamma'(\phi(t)) \sigma''(t)$ implies that $\sigma''(t)$ lies in the plane spanned by $T(\phi(t))$ and $N(\phi(t))$. (This plane is called the osculating plane.) In practical calculations in 3-space one calculates $\sigma''$, then calculates $B$ by normalizing the cross-product $\sigma' \times \sigma''$. Only afterwards(!) one calculates $N = B \times T$. Returning to the acceleration $\sigma''$, the magnitudes of its tangential and normal components are easily calculated as

$$a_\parallel(t) = T(\phi(t)) \cdot \sigma''(t) = \frac{\sigma'(t) \cdot \sigma''(t)}{\|\sigma'(t)\|} \quad \text{and} \quad a_\perp(t) = N(\phi(t)) \cdot \sigma''(t) = \pm \sqrt{\|\sigma''(t)\|^2 - a_\parallel^2(t)} \quad (6)$$

The curvature (and radius of curvature $\rho = \frac{1}{\kappa}$) and the torsion may be obtained in various ways. Typical formulas suitable for practical calculations (for space curves) are

$$\kappa(\phi(t)) = \frac{|a_\perp(t)|}{\|\sigma'(t)\|^2} = \frac{\|\sigma'(t) \times \sigma''(t)\|}{\|\sigma'(t)\|^3} \quad \text{and} \quad \tau(\phi(t)) = \frac{\sigma'(t) \times \sigma''(t) \cdot \sigma'''(t)}{\|\sigma'(t) \times \sigma''(t)\|^2} \quad (7)$$

Exercise 1.14 Derive the formulas presented above for $\kappa(\phi(t))$ and for $\tau(\phi(t))$ from the definitions of $\kappa$ and $\tau$ in terms of the Frenet formulas.

Exercise 1.15 Consider smooth planar curves $\sigma: J \mapsto \mathbb{R}^2$. Devise a practical strategy to calculate $T, N, \kappa$ with minimal effort. (Note, that from $T$ one easily obtains $N$ by interchanging the components and changing the sign of one of the components. Which one? Why?)

Exercise 1.16 Derive the usual formula $\kappa = y''/\left(1 + (y')^2\right)^{3/2}$ for the curvature in the special case of graphs of functions, i.e. curves of the form $\gamma(t) = (t, f(t))$.

Exercise 1.17 Verify that the curve $\gamma: \mathbb{R} \mapsto \mathbb{R}^2$ defined by $\gamma(t) = (\exp(-1/t^2), 0)$ if $t > 0$, $\gamma(0) = (0, 0)$ and $\gamma(t) = (0, \exp(-1/t^2), 0)$ if $t < 0$ is infinitely many times continuously differentiable on the whole real line. Describe the $(T(s), N(s))$-frame for this curve (this is very easy), with special attention to what happens when $t = 0$. 
Project 1.18 Write a MAPLE procedure that takes as input a space curve, i.e. algebraic expressions for $x(t), y(t)$ and $z(t)$ (and possibly other parameters such as the domain $I$), and which gives as output an animation of the Frenet frame along the curve. As test curves consider a helix $\sigma(t) = (\cos t, \sin t, t)$ and the $(2,3)$-torus-knot $\sigma = u^{-1} \circ \gamma$ where $\gamma(t) = (2t, 3t)$ for $t \in [0, 2\pi]$ and $u^{-1}(\theta, \phi) = ((R + r \sin \phi) \cos \theta, (R + r \sin \phi) \sin \theta, r \cos \phi)$, e.g. for $R = 5$ and $r = 2$.

Project 1.19 Consider a family of closed curves $\gamma(s, t)$ that are parameterized by arc-length $s$ and that evolve with time $t$ according to e.g. the heat-equation $\frac{\partial^2}{\partial s^2} \kappa(s, t) - \frac{\partial}{\partial t} \kappa(s, t) = 0$. Intuitively, this is the easiest model that describes how loops may try to straighten out under the influence of tension. For simplicity, start with initial curvature functions $\kappa(s, 0)$ that are expressed as (finite) Fourier polynomials $\kappa(s, 0) = a_0 + \sum_{j=1}^{N} (a_j \cos jt + b_j \sin jt)$. This allows one to explicitly write out the solutions $\kappa(s, t) = a_0 + \sum_{j=1}^{N} e^{-j^2t}(a_j \cos jt + b_j \sin jt)$ of the heat equation. First find conditions on the Fourier coefficients that assure that the associated curve is closed. Then integrate the two-dimensional analogue of the Frenet-Serret formulas to obtain the associated curve $\gamma(t, s)$. Animate the images of the curves.

An exploratory worksheet that addresses this project is available from the WWW-site http://math.asu.edu/~kawski/MAPLE/MAPLE.html. However, it has at least a cosmetic flaw as it arbitrarily fixes $\gamma(0, t)$ and $\frac{\partial}{\partial s} \gamma(0, t)$ – a nicer solution would include a more physical solution, e.g. fixing the center of mass of the curve by translating the curve as needed. Even better, it would be nice to add, if necessary, a rotation so that the total angular momentum as determined from $\frac{\partial}{\partial t} \gamma(t, 0)$ is constant equal to zero. Any improvements of this worksheet are highly welcome. They likely will lead to further, even more interesting applications – starting with an analogous exploration of loops in 3-space!

In a future version add a little classical stuff involving evolutes and involutes – much of this can be done in exercises. Alternatively, do this in some MAPLE worksheets. Till then refer to Oprea’s book as a nice reference.
2 Manifolds

2.1 Introduction

We want to think of manifolds as abstractions and generalizations of the intuitive notions of curves and surfaces. This subsection reviews a few key ideas, purposes and examples. The next subsection provides a few fundamental topological notions to prepare for a precise definition of manifolds, first in the topological category, and then in the differential category.

The upcoming definition will characterize a manifold as a space which is such that every point in it has a neighbourhood that is homeomorphic to an open subset of a Euclidean space $\mathbb{R}^n$. In particular, we shall not allow for edges and boundaries to avoid the associated technical complications. Next, we will equip manifolds with differentiable structures that allow for notions such as dynamical systems evolving on the manifolds, and for generalized notions of curvature. Typical objectives are to analyze the effects of curvature on the global topological structure or on the behaviour of dynamical systems. A need to integrate over (subsets of) manifolds arises naturally. A major role of local coordinate charts is to transfer these differential (and integral) concepts back into familiar Euclidean space where standard techniques may be employed for calculations.

Throughout we will emphasize geometric points of view – as a simple example what we don’t want think of the two dimensional sphere $S^2$ as (the union of) the graph(s) of two functions $z = \pm \sqrt{x^2 + y^2}$. This rather arbitrary preferential treatment of $z$ versus $x$ and $y$ begins to hide the full symmetry of the sphere under a group of rotations and reflections.

Before proceeding to technical descriptions let us take a brief look at some typical examples that should be included in our notion of manifold.

Curves and surfaces, especially the Euclidean spaces $\mathbb{R}^n$, and (open) subsets of Euclidean spaces should be manifolds. However, we may impose conditions so as to avoid e.g. self-intersections, boundaries, and, in the category of differentiable manifolds, cusps, corners and the like. The characterization of the two dimensional sphere $S^2 \subseteq \mathbb{R}^3$ as the set of all $(x, y, z) \in \mathbb{R}^3$ that satisfy $x^2 + y^2 + z^2 = 1$, invites a natural generalization to higher dimensional analogues of surfaces as subsets of $\mathbb{R}$ that may be characterized by (sets of) equations $F_k(x_1, x_2, \ldots x_n) = 0$ ($k = 1, \ldots p$). To avoid cusps and corners one usually imposes a condition that the gradient (or a higher dimensional analogue) does not vanish.

As a special case, this description immediately opens the door to objects such as the group of special orthogonal matrices $SO(3)$ in three or $n$ dimensions. The defining equation $A^\top A = I_{n\times n}$ is of the same form as the equation of the sphere given above. What makes these matrix manifolds particularly interesting is their natural group structure – there is a natural notion of multiplying points on the generalized surface – this is the starting point for Lie groups.

A different way that many manifolds of interest are obtained is by taking quotients. In the most simple case the circle $S^1$ arises as a quotient of $\mathbb{R}$ by $\mathbb{Z}$. Intuitively, for any periodic function $f$ with period $p > 0$, i.e. $f(x+p) = f(x)$ for all $x \in \mathbb{R}$, one may consider as its natural domain any interval $[a, a+p]$ with endpoints identified. More abstractly, consider the equivalence relation $\sim$ defined on $\mathbb{R}$ by $x \sim y \iff (x - y)/p \in \mathbb{Z}$. Then each point on the circle $\Theta$ represents an equivalence class $[\Theta] = \{\Theta + kp : k \in \mathbb{Z}\}$.

In an analogous way, the torus arises naturally (e.g. very commonly in dynamical systems) as the quotient of the plane $\mathbb{R}^2$ by $\mathbb{Z}^2$. One commonly visualizes the torus as the unit square $[0, 1] \times [0, 1]$ with opposing edges identified.
If one starts with the same square, but identifies one (or two) sets of opposing edges with orientation reversed one arrives at the Klein bottle and at the projective plane. Neither one of these can be visualized in the usual way as a surface in $\mathbb{R}^3$, but apparently each shares many properties with the torus due to their analogous construction.

More abstractly, projective spaces arise when considering the spaces of all (straight) lines in $\mathbb{R}^n$ that pass through the origin. Before looking at this more closely, recall the simple case of considering the space of all (semi-infinite, open) rays emanating from the origin. Each of these rays may be naturally identified with the point on the unit sphere (unit circle) through which it passes. Thus we may think of the spheres $S^{n-1}$ as arising from $\mathbb{R}^n \setminus \{0\}$ as quotients under the equivalence relation $x \sim y \iff$ if there exists $\lambda \in \mathbb{R}$, $\lambda > 0$ such that $x = \lambda y$. In a practical sense this is closely related to considering only the angle(a) $\theta$ (or $(\theta, \phi)$) when working with polar (or spherical coordinates).

In analogy, if one discards the requirement $\lambda > 0$ in the preceding definition of equivalence, then the equivalence classes are the lines through the origin. (More precisely, since we started with $\mathbb{R}^n \setminus \{0\}$, the origin is removed from each line, or the line really consists of two rays.) One may visualize the resulting quotient space as the space of pairs of opposite points on the sphere $S^{n-1}$, or as a semi-sphere with two halves of the equator (which is a sphere $S^{n-2}$ by itself) identified, or glued together with careful attention to the orientation of each piece.

From these projective spaces it is only a small step to Grassmannian manifolds which may be thought of as spaces of $m$-dimensional (hyper-)planes in $n$-dimensional Euclidean space.

A typical application where these appear naturally is in the classification of linear control systems $\dot{x} = Ax + Bu$, with state $x \in \mathbb{R}^n$ and control $u \in \mathbb{R}^m$. Here one considers two systems equivalent if one may be transformed into the other by coordinates changes, in state $\tilde{x} = Rx$ and control space $\tilde{u} = Su$, and/or under feedback transformations $\tilde{u} = u + Kx$, . . .

All the above examples clearly have (preserve) some additional structure beyond just being sets of points. In order to be able to work with concepts such as continuity and notions of derivatives one intuitively needs some notion of distance. Indeed, while one can start with even more general topological spaces, in the finite dimensional setting very little is lost if one requires that the set is equipped with at least some a-priori notion of distance. However, this basic notion of distance will primarily be used only as a foundation for e.g. continuity, and should not be confused with the Riemannian metrics that we will study later, and which have a deeply connected with curvature.
2.2 Some basic topological notions

This subsection reviews some basic definitions and properties of objects in topology.

Definition 2.1 A metric on a set $X$ is a function $d : X \times X \to \mathbb{R}$ that satisfies

(i) $d(x, y) \geq 0$ for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$ (positive definiteness),

(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$ (symmetry), and

(iii) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$ (triangle inequality).

A metric space is a pair $(X, d)$ where $X$ is a set and $d$ is a metric of $X$.

A set $X$ may be equipped with different metrics, which in general result in completely different notions of continuity. For example, common metrics on $\mathbb{R}^n$ are

- the usual Euclidean distance defined by $d_2(x, y) = \|y - x\| = \sqrt{(y_1 - x_1)^2 + \ldots + (y_n - x_n)^2}$
- the taxi-cab metric defined by $d_1(x, y) = \sum_i |y_i - x_i|$, and
- the sup-norm $d_\infty(x, y) = \max_i |y_i - x_i|$.

Two metrics $d$ and $d'$ on a space $X$ are called equivalent if there exists constants $c, C > 0$ such that $cd(x, y) \leq d(x, y) \leq Cd'(x, y)$ for all $x, y \in X$.

Exercise 2.1 Show that the three metrics on $\mathbb{R}^n$ discussed above are equivalent. Show pictorially the meaning of above inequalities in terms of nested $\varepsilon$-balls with respect to the different metrics. Conclude that the open sets in $\mathbb{R}^n$ are the same, independent of the metric employed to define open balls.

On any set $X$ the discrete metric may be defined by $d(x, y) = 1$ is $x \neq y$ and $d(x, y) = 0$ if $x = y$.

Given a metric $d$ on a space $X$ one may construct from it a bounded metric $\bar{d}$ by setting $\bar{d}(x, y) = d(x, y)$ if $d(x, y) \leq 1$ and $\bar{d}(x, y) = 1$ else. Another useful bounded metric may be obtained by defining $\tilde{d} = d/(1 + d)$.

Exercise 2.2 Verify that the discrete and bounded metric described above are indeed metrics.

Exercise 2.3 Let $X$ be the set of all lines in the plane that pass through the origin. For lines $\ell_1$ and $\ell_2$ let $d(\ell_1, \ell_2)$ to be the (smaller) angle between them. Verify that $d$ is a metric on $X$.

Definition 2.2 Suppose $(X, d)$ is a metric space. A set $O \subseteq X$ is called open if for every $p \in O$ there is an $\varepsilon > 0$ such that the (open) $\varepsilon$-ball $B_p(\varepsilon) = \{x \in X : d(x, p) < \varepsilon\}$ is contained in $O$, i.e. $B_p(\varepsilon) \subseteq O$. A set $F \subseteq X$ is called closed if $X \setminus F$ is open.

Exercise 2.4 Show that if $d$ is the discrete metric on a set $X$ then every subset $S \subseteq X$ is both open and closed.

Exercise 2.5 Consider a metric space $(X, d)$ and the bounded metric $\bar{d}$ (or $\tilde{d}$) constructed from $d$ as above. Show that a subset $S \subseteq X$ is open in $(X, d)$ if and only if it is open in $(X, \bar{d})$. 
The notions of open and closed do not require an underlying metric structure. The following axioms allow for a generalization to spaces without a metric:

**Definition 2.3** A topology on a set $X$ is a collection $\mathcal{T} \subseteq 2^X$ of subsets of $X$ that satisfies

1. $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,
2. $\mathcal{T}$ is closed under (arbitrary) unions, i.e., if $\{O_\alpha : \alpha \in A\} \subseteq \mathcal{T}$ then $\bigcup_{\alpha \in A} O_\alpha \in \mathcal{T}$, and
3. $\mathcal{T}$ is closed under finite intersections, i.e., if $O_k \in \mathcal{T}$, $k = 1, 2, \ldots, n$, then $\bigcap_{k=1}^n O_k \in \mathcal{T}$.

A subset $O \subseteq X$ is called open if $O \in \mathcal{T}$. A subset $F \subseteq X$ is called closed if $X \setminus F \in \mathcal{T}$.

A topological space is a pair $(X, \mathcal{T})$ where $X$ is a set and $\mathcal{T}$ is a topology on $X$.

Note that an infinite intersection of open sets is not required to be open. The standard example is the real line with the usual topology and $O_k = (-\frac{1}{k}, \frac{1}{k})$. Clearly $\bigcap_{k=1}^\infty O_k = \{0\}$ which is not open (in the standard topology).

The commonly used open neighbourhood of $p \in X$ refers to an open set which contains $p$.

While technically a topological space is a pair $(X, \mathcal{T})$, one often refers to $X$ alone as a topological space. In such cases it is usually understood from the context which topology $\mathcal{T}$ on $X$ is meant.

**Usual to specify only basis or subbasis for a topology. 1st/2nd countable. Defn. Exa. Exer.**

**Exercise 2.6** Show that if $Y \subseteq X$ is a subset of a topological space $(X, \mathcal{T})$ then the collection $\mathcal{T}' = \{O \cap Y : O \in \mathcal{T}\}$ defines a topology on $Y$. This topology is called the subspace topology.

**Exercise 2.7** Suppose that $(X, d)$ is a metric space. Show that the collection of open sets (in the sense of open in a metric space) defines a topology on $X$. (This topology is called the metric topology on $(X, d)$.)

**Definition 2.4** Suppose $X$ and $Y$ (or, more precisely $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$) are topological spaces.

A map $f : X \to Y$ is called continuous if for every open set $O \subseteq Y$ the preimage $f^{-1}(O) \subseteq X$ is open (i.e. $O \in \mathcal{T}_Y \implies f^{-1}(O) \in \mathcal{T}_X$). (This is equivalent to $f^{-1}(F) \subseteq X$ closed for every $F \subseteq Y$ closed.)

A map $f : X \to Y$ is called open if for every open set $O \subseteq X$ the image $f(O) \subseteq Y$ is open.

A map $f : X \to Y$ is called closed if for every closed set $F \subseteq X$ the image $f(F) \subseteq Y$ is closed.

In the case that $\mathcal{T}_X$ and $\mathcal{T}_Y$ are the metric topologies associated with metrics $d_X$ and $d_Y$ on $X$ and $Y$, respectively, this notion of continuity agrees with the standard $\varepsilon$-$\delta$ characterization of continuity. A function $f : X \to Y$ is continuous (as defined above) if and only if for every $p \in X$ and for every $\varepsilon > 0$ there is a $\delta > 0$ such that if $q \in X$ with $d_X(p, q) < \delta$ then $d_Y(f(p), f(q)) < \delta$.

Practically the notion of continuity captures the concept that small changes in the input of a function cause only small changes in the output.

**Exercise 2.8** Consider the set $\mathbb{R}$ of real numbers with the usual topology $\mathcal{T}_2$, with the indiscrete topology $\mathcal{T}_1 = \{\emptyset, \mathbb{R}\}$, and with the discrete topology $\mathcal{T}_3$ in which every subset of $\mathbb{R}$ is open.

For each pair $(i, j)$ with $i, j = 1, 2, 3$ describe the set of continuous functions from $(\mathbb{R}, \mathcal{T}_i)$ to $(\mathbb{R}, \mathcal{T}_j)$. (Make a $3 \times 3$ table.) In particular, for which pairs is the identity function $id : x \mapsto x$ continuous? For which pair(s) are (only) the constant functions continuous, and for which pair(s) are all functions continuous?

**Exercise 2.9** Verify the assertion that in metric spaces the standard $\varepsilon$-$\delta$ characterization of continuity agrees with the definition given above.
Definition 2.5
A map \( f : X \to Y \) between topological spaces \( X \) and \( Y \) is called a homeomorphism if

(i) \( f \) is a bijection, i.e. one-to-one and onto,

(ii) \( f \) is continuous, and

(iii) \( f^{-1} \) is continuous (i.e. \( f \) is open).

Two spaces topological spaces \( X \) and \( Y \) are called homeomorphic if there exists a homeomorphism \( f \) that maps \( X \) onto \( Y \).

Do not confuse the term homeomorphism discussed here with homomorphism which refers to maps that preserve algebraic relationships as in \( f(p \cdot q) = f(p) \cdot f(q) \).

From a topological point of view homeomorphic spaces are basically considered as identical.

Exercise 2.10 Verify that the map \( f : (0,1) \to \mathbb{R} \) defined by \( f(x) = (1-2x)/(x(x-1)) \) is a homeomorphism. (Calculate \( f' \) and consider \( \lim_{x \to 0^+} f(x) \) and \( \lim_{x \to 1^-} f(x) \).)

Exercise 2.11 Verify that the map \( f : \mathbb{R}^2 \to B^2_0(1) = \{ x \in \mathbb{R}^2 : \| x \| < 1 \} \) defined by \( f(x) = x/(1+\| x \|) \) is a homeomorphism.

While it appears intuitively clear that \( \mathbb{R}^n \) and \( \mathbb{R}^m \) are not homeomorphic for \( n \neq m \), the proof (for general \( m \) and \( n \)) is surprisingly hard – usually utilizing tools from algebraic topology.

Note that there exists for example a continuous function \( f : [0,1] \to [0,1] \times [0,1] \) that is both one-to-one and onto. However, the inverse \( f^{-1} \) can’t be continuous. For more details on the construction of such space-filling curves see e.g. Munkres, Topology, p. 271.

Definition 2.6 A subset \( A \subseteq X \) of a topologically space \( X \) is called connected if whenever \( B, C \subseteq X \) are disjoint open sets such that \( B \cup C = A \) then \( A \subseteq B \) or \( A \subseteq C \) (i.e. \( A \cap B = \emptyset \) or \( A \cap C = \emptyset \)). If \( A \) is not connected then it is called disconnected.

Exercise 2.12 Suppose that \( f : X \to Y \) is a continuous map. Show that if \( f \) is onto and \( X \) is connected then \( Y \) is connected.

Exercise 2.13 On a topological space \( X \) define the relation \( \sim \subseteq X \times X \) by \( x \sim y \) if there exists a connected subset \( C \subseteq X \) such that \( x \in C \) and \( y \in C \). Show that \( \sim \) is an equivalence relation. (The equivalence classes of this relation are called the (connected) components of the space \( X \).)

In general topological spaces one works with a number of different notions of connectedness. Here we only mention the following other notion, which is stronger than connectedness:

Definition 2.7 A subset \( A \subseteq X \) of a topologically space \( X \) is called path-connected if whenever \( p, q \in A \) then there exist a continuous map \( \gamma : [0,1] \to A \) such that \( \gamma(0) = p \) and \( \gamma(1) = q \).

Exercise 2.14 Show that every path-connected set is connected.

Exercise 2.15 Show that the topologist’s sine curve, i.e. the set \( \{ (x, \sin(\frac{1}{x})) : x \neq 0 \} \cup \{ 0 \} \times [-1,1] \) is connected but not path-connected.
One of the most common uses of connectedness is the argument that if a function \( f: X \to \mathbb{R} \) is continuous and locally constant then it is constant provided the domain \( X \) is connected. Here, **locally constant** means that every \( p \in X \) has an open neighbourhood \( U \) (containing \( p \)) such that the restriction of \( f \) to \( U \) is constant.

To clarify this argument, consider the function \( f: \mathbb{R} \setminus \{0\} \to \mathbb{R} \) defined by \( f: x \mapsto 0 \) if \( x < 0 \) and \( f: x \mapsto 1 \) if \( x > 0 \). Clearly the derivative \( f' \equiv 0 \) vanishes identically, but \( f \) is not constant. Of course, the key is that the domain is not connected. Consequently, the vanishing of the derivative only assures that \( f \) is locally constant. It does not assure that \( f \) is constant.

Arguably the most important topological concept for us is compactness. One may think of it as an outgrowth of the desire to generalize, or to get to the real basics of the important theorem that every continuous function \( f: [a, b] \to \mathbb{R} \) defined on a closed bounded interval attains its minimum and its maximum, i.e. there exist points \( x_1, x_2 \in [a, b] \) such that \( f(x_1) \leq f(x) \leq f(x_2) \) for all \( x \in [a, b] \). It is well-known from freshmen calculus that this assertion fails if either of the closedness or boundedness hypotheses is omitted. The closedness requirement naturally generalizes to general topological spaces, but the boundedness does not. For example, if \( \mathbb{R}^n \) is equipped with the bounded metric \( \bar{d} = d_2/(1 + d_2) \) (where \( d_2 \) is the standard Euclidean metric), then \( (\mathbb{R}^n, \bar{d}) \) still has the same topology as \( (\mathbb{R}^n, d_2) \) yet while \( K = \mathbb{R}^n \) is closed and bounded in \( (\mathbb{R}^n, d) \), it is not in \( (\mathbb{R}^n, \bar{d}) \). Many different generalizations have been proposed to generalize the basic idea of “closed and bounded” which is so useful in \( \mathbb{R}^n \) (with its usual metric). Any introductory course in point-set topology will discuss such different notions of compactness. It was not until quite late into the 20th century that the following notion finally crystallized, and it became clear that it captures the fundamental features of the desired properties.

**Definition 2.8** A subset \( K \subseteq X \) of a topological space \( X \) is called compact if every open cover of \( K \) has a finite subcover, i.e. if \( \{O_\alpha \subseteq X: \alpha \in A\} \) is a collection of open sets such that \( K \subseteq \bigcup_{\alpha \in A} O_\alpha \), then there exists a finite subcollection \( \{O_{\alpha_j}: j = 1, 2, \ldots, n\} \) such that \( K \subseteq \bigcup_{j=1}^n O_{\alpha_j} \).

The Heine-Borel theorem asserts that every bounded closed interval in \( \mathbb{R} \) is compact. Its proof may be found in any advanced calculus text.

**Exercise 2.16** Prove that if \( f: X \to Y \) is continuous and \( K \subseteq X \) is compact then the image \( f(K) \subseteq Y \) is compact.

In the case of \( Y = \mathbb{R} \) this implies that there exist points \( p, q \in X \) at which \( f \) attains its global minimum and global maximum, i.e. such that \( f(p) \leq f(x) \leq f(q) \) for all \( x \in X \).

**Definition 2.9** A sequence \( \{a_k\}_{k \in \mathbb{N}} \subseteq X \) in a topological space \( X \) is said to converge if there exists \( \bar{x} \in X \) such that for every open set \( O \subseteq K \) containing \( \bar{x} \), there exist a finite natural number \( N \) such that \( a_n \in O \) for all \( n > N \).

**Exercise 2.17** Suppose \( \{x_k\}_{k \in \mathbb{N}} \) is an infinite sequence with values in a compact topologically space \( K \). Show that \( \{x_k\}_{k \in \mathbb{N}} \) has an accumulation point \( \bar{x} \in K \). If, in addition, \( K \) is first countable then there exists a converging subsequence \( \{x_{k_j}\}_{j \in \mathbb{N}} \).
Finally we mention a few separation axioms which on occasion are used as essential hypotheses in differential geometry.

**Definition 2.10**

- A topological space $X$ is called a Hausdorff space if for every pair of distinct points $p, q \in X$ there exist disjoint open sets $U$ and $V$ such that $p \in U$ and $q \in V$.

- A Hausdorff space $X$ is called completely regular if one-point sets are closed in $X$ and if for every point $p \in X$ and every closed set $F \subseteq X$ not containing $p$ there exists a continuous function $f: X \rightarrow \mathbb{R}$ such that $f(p) = 0$ and $f(x) = 1$ for every $x \in F$.

- A Hausdorff space $X$ is called normal if one-point sets are closed in $X$ and if for every pair of disjoint closed sets $F_1, F_2 \subseteq X$ there exists disjoint open sets $U$ and $V$ such that $F_1 \subseteq U$ and $F_2 \subseteq V$.

These notions become useful when patching together local results, e.g. obtained in one coordinate chart at a time. This will be made precise when discussing partitions of unity in a subsequent section.

For the sake of completeness we here also give the following two technical definitions:

**Definition 2.11** A map $f: X \rightarrow Y$ is called proper if for every compact set $K \subseteq Y$ the preimage $f^{-1}(K) \subseteq X$ is compact.

**Definition 2.12** A subset $S \subseteq X$ of a Hausdorff space $X$ is called paracompact if every open cover of $S$ has a locally finite open refinement. This means that if $U_\alpha, \alpha \in A$ are open sets such that $S \subseteq \bigcup_{\alpha \in A} U_\alpha$ then there exist a collection of open sets $V_\beta, \beta \in B$ such that

- For every $\beta \in B$ there exists an $\alpha \in A$ such that $V_\beta \subseteq U_\alpha$,

- $S \subseteq \bigcup_{\beta \in A} V_\beta$, and

- every $p \in S$ has an open neighbourhood $W$ which intersects only a finite number of the sets $V_\beta, \beta \in B$.

Paracompactness is very close to metrizability, (indeed, metrizability is equivalent to paracompactness and local metrizability). Thus many authors use paracompactness as a basic requirement when defining manifolds.
2.3 Local coordinate charts

This section defines the concept of a topological manifold which is to serve as a spring board for the subsequent definition of a differentiable manifold. The main focus is on the concept of local coordinate charts.

**Definition 2.13** A topological manifold $M$ is a metric space $(M,d)$ such that for every $p \in M$ there exist an open set $U \subseteq M$ containing $p$ and a homeomorphism $u: U \mapsto \mathbb{R}^n$ for some $n \in \mathbb{N}$.

A few remarks:

- The metric $d$ plays little role in the future – what is needed is really only a reasonably nice topological space. Metric, or more accurately, metrizable spaces just happen to have about the right set of properties needed throughout the standard development.

- The statement that “$U$ is homeomorphic to $\mathbb{R}^n$” may be replaced by “$U$ is homeomorphic to an open subset of $\mathbb{R}^n$.”

- If $M$ is connected, then $n$ is constant and is called the dimension of the manifold $M$.

- The functions $u^i = x^i \circ u: U \mapsto \mathbb{R}$ are called local coordinates. We use $x^i: \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \ldots n$ to denote the standard coordinate functions, defined by $x^i: (a_1, \ldots a_n) \mapsto a_i$ for every $a = (a_1, \ldots a_n) \in \mathbb{R}^n$. The pair $(u, U)$ is called a local coordinate chart about $p$. Note that if $(u, U)$ is a local coordinate chart about $p$ then $(\tilde{u}, U)$ defined by $\tilde{u}: q \mapsto u(q) - u(p)$ for $q \in U$ is a local coordinate chart about $p$ such that $\tilde{u}(p) = 0$. This is usually written as $\tilde{u}: (U, p) \mapsto (\mathbb{R}^n, 0)$.

![Picture: schematic U, V, u(U), v(V) and maps]

If $(u, U)$ and $v, V$ are coordinate charts about $p \in M$ (i.e. in particular $p \in U \cap V$, then

$$v \circ u^{-1}: u(U \cap V) \mapsto v(U \cap V)$$

is a homeomorphism between open subsets of $\mathbb{R}^n$.

We continue with a short list of examples of manifolds and coordinate charts.

- For any $n \geq 0$ the $n$-dimensional Euclidean space $\mathbb{R}^n$ is an $n$-dimensional manifold with the single coordinate chart $(\text{id}, \mathbb{R}^n)$.

- The open ball $B_p(r) = \{ x \in \mathbb{R}^n: \|x - p\| < r \}$ of radius $r$ about $p \in \mathbb{R}^n$ is an $n$-dimensional manifold with a single chart given by $U = B_p(r)$ and $u(x) = (x - p)/(r - \|x - p\|)$. 
Exercise 2.18 Verify that the inverse is given by \( u^{-1}(y) = p + ry/(1 + \|y\|) \).

- Every open subset of an \( n \)-dimensional manifold is itself an \( n \)-dimensional manifold.
- Identify the space \( M_{m,n}(\mathbb{R}) \) of \( m \)-by-\( n \) matrices with real entries with the space \( \mathbb{R}^{mn} \).

E.g. in the \( 2 \times 2 \)-case simply identify

\[
u: \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}
\]

Note that with this identification \( M_{m,n}(\mathbb{R}) \) inherits a natural metric structure. Clearly this shows that the entire space \( M_{m,n}(\mathbb{R}) \) is an \( nm \)-dimensional manifold.

Of more interest are various subspaces of \( M_{m,n}(\mathbb{R}) \). Typical examples are the general linear group \( GL(n, \mathbb{R}) = \{ A \in M_{n,n}(\mathbb{R}) : \det A \neq 0 \} \) and its subset of orientation preserving nonsingular matrices \( GL^+(n, \mathbb{R}) = \{ A \in M_{n,n}(\mathbb{R}) : \det A > 0 \} \). Both are \( n^2 \)-dimensional manifolds. The argument uses that \( \det \) is a polynomial function in the entries of the matrix, and hence it is continuous. Consequently, the preimages \( \det^{-1}(\mathbb{R} \setminus \{0\}) \) and \( \det^{-1}(0, \infty) \) of open sets are open subsets of \( M_{m,n}(\mathbb{R}) \).

Further examples are the special linear groups \( SL(n, \mathbb{R}) = \{ A \in M_{n,n}(\mathbb{R}) : \det A = 1 \} \), the orthogonal groups \( O(n) = \{ A \in M_{n,n}(\mathbb{R}) : A^\top A = I \} \), and the special orthogonal groups \( SO(n) = O(n) \cap SL(n, \mathbb{R}) \). Unlike the previous examples – which are open subsets – these manifolds are defined by closed conditions (i.e. "=" as opposed to "\(!=\)", "<" or "\(>\)"). We return to these examples later when tools from differential calculus on manifolds will make it easy to establish when such subsets defined by closed conditions give rise to manifolds.

- The \( m \)-sphere \( S^m(r) = \{ x \in \mathbb{R}^{m+1} : \|x\| = r \} \) is a \( m \)-manifold. The case \( m = 0 \) is special with \( S^0(r) = \{-r, r\} \subseteq \mathbb{R} \) consisting of only two points, and thus being disconnected. The 1-sphere is a circle, and one needs at least two charts, e.g. \( U = S^1(r) \setminus \{(r, 0)\} \) and \( V = S^1(r) \setminus \{(-r, 0)\} \). Use \( u(x, y) = \arctan(2x, y) \) taking values in \( (-\pi, \pi) \) and \( v(x, y) = \arctan(2x, y) \) taking values in \( (0, 2\pi) \). Note that the local coordinates agree essentially with the angle of polar coordinates.

Exercise 2.19 Construct a collection of charts for the 2-sphere \( S^2(r) \) by explicitly adapting the spherical coordinates \( u(x, y, z) = (\theta(x, y, z), \phi(x, y, z)) \in (0, 2\pi) \times (0, \pi) \) to different subsets resulting from different "cuts". What are the minimal number of cuts, and the minimal number of charts needed to cover \( S^2 \)?

A different set of coordinates, that is particularly useful for higher dimensional spheres is based on stereographic projections: Consider the two subsets \( U_\pm = S^m(r) \setminus \{(0, 0, \ldots, \pm r)\} \) and define the maps \( u_\pm: U_\pm \mapsto \mathbb{R}^m \) by

\[
u_\pm(x) = \frac{2r}{2r \mp x_{m+1}}(x_1, \ldots, x_m)
\]

Graphically \( (u_\pm(x), \mp r) \in \mathbb{R}^{m+1} \) is the point where the hyperplane \( x_{m+1} = \mp r \) intersects the line that passes through the point \( x \in S^m(r) \) and through \( (0, 0, \ldots, \pm r) \).
Exercise 2.20 For the inverse maps $u^{-1}_\pm: \mathbb{R}^m \mapsto U_\pm \subseteq S^m(r)$, derive the formulae

$$x = u^{-1}_\pm(y) = \left(\frac{2r}{1 + \|\frac{y}{2r}\|^2} \left(\frac{y_1}{2r}, \ldots, \frac{y_m}{2r}\right), \frac{1 - \|\frac{y}{2r}\|^2}{1 + \|\frac{y}{2r}\|^2}\right)$$

(10)

Use this to obtain explicit formulae for the “transition maps” $u_\pm \circ u^{-1}_\pm: \mathbb{R}^m \mapsto \mathbb{R}^m$.

What do these maps do graphically e.g. which sets do they leave fixed?

What are the images of (special) lines and circles?

- If $M^m$ and $N^n$ are $m$- and $n$-dimensional manifolds, respectively, then the Cartesian product $M \times N$ is an $(m+n)$-dimensional manifold: Suppose $(p, q) \in M \times N$ and $(u, U)$ and $(v, V)$ are coordinate charts about $p$ and $q$, then $(u \times v, U \times V)$ is a coordinate chart about $(p, q)$ where $(u, v)(a, b) = (u(a), v(b)) \in \mathbb{R}^m \times \mathbb{R}^n$.

A typical example uses that the circle $S^1 = \{x \in \mathbb{R}^2: ||x|| = 1\}$ is a manifold to establish that the torus $T^2 = S^1 \times S^1$ is a 2-dimensional manifold.

- We briefly return to the real projective spaces, now illustrating coordinate charts. On $\mathbb{R}^{m+1} \setminus \{0\}$ define the equivalence relation $x \sim y$ if $x = \lambda y$ for some $\lambda \in \mathbb{R}$. The $m$-dimensional real projective space is defined as the quotient $\mathbb{P}^m = \left(\mathbb{R}^{m+1} \setminus \{0\}\right) / \sim$, i.e. a point $[x] \in \mathbb{P}^m$ is the equivalence class $[x] = \{y \in \mathbb{R}^{m+1} \setminus \{0\}: y \sim x\}$. Intuitively think of $\mathbb{P}^m$ as the space of all lines in $\mathbb{R}^{m+1}$ that pass through the origin, or as the $m$-sphere $S^m$ with antipodal points $x$ and $-x$ identified. More graphically, one may obtain $\mathbb{P}^2$ by sewing a disk to the (only one!) edge of a Möbius strip.

For $j = 1, \ldots, m$ consider the sets $U_j = \{[x] \in \mathbb{P}^m: x_j \neq 0\}$ and coordinates maps (homogeneous coordinates)

$$u_j([x]) = \left(\frac{x_1}{x_j}, \ldots, \frac{x_{j-1}}{x_j}, \frac{x_{j+1}}{x_j}, \ldots, \frac{x_{m+1}}{x_j}\right)$$

(11)

It is straightforward to verify that the value of $u_j([x])$ does not depend on the choice of the representative $x \in [x]$. The inverse is given by $u^{-1}_j(y_1, \ldots, y_m) = [y_1, \ldots, y_{i-1}, 1, y_{i+1}, \ldots, y_m]$.

A complete discussion of these coordinate maps (they are supposed to be homeomorphisms) is straightforward but technical, in terms of the quotient topology. For an introductory discussion of quotient maps, and quotient topologies see e.g. Munkres (Topology, a first course”, p.134). A main issue is to assure that the quotient topology is not pathological.

Just as a reference, a surjective map $f: X \mapsto Y$ is called a quotient map if $O \subseteq Y$ is open if and only if $f^{-1}(O) \subseteq X$ is open. For any map $f: X \mapsto A$ from a topological space $X$ to a set $A$ there is exactly one topology on $A$, called the quotient topology, such that $f$ is a quotient map. In the case that $A$ is a set of equivalence classes on $X$, $A$ with this topology is called a quotient space of $X$.

- Two-dimensional surfaces in $\mathbb{R}^3$, or more generally $m$-dimensional hypersurfaces in $\mathbb{R}^{m+1}$ are some of the most commonly used manifolds. Clearly every graph $\{(x, f(x)): x \in \mathbb{R}^m\} \subseteq \mathbb{R}^{m+1}$ of any continuous function $f: \mathbb{R}^m \mapsto \mathbb{R}$ is a manifold with a single chart $u: (x, f(x)) \mapsto x \in \mathbb{R}^m$.

More interesting are hypersurfaces that arise as preimages $F^{-1}(\{0\})$ ("zero-sets") $M = \{x \in \mathbb{R}^{m+1}: F(x) = 0\}$ of functions $F: \mathbb{R}^{m+1} \mapsto \mathbb{R}$, or that are given by parameterizations
More generally, if $M$ and $N$ are manifolds and $\Phi: M \rightarrow N$ then $\Phi(M) \subseteq N$ may be a manifold. Similarly, if $P \subseteq N$ is a submanifold, then $\Phi^{-1}(P) \subseteq M$ might be a submanifold of $M$. To avoid unnecessary duplication and difficulties we shall discuss these constructions only in the setting of \emph{differentiable manifolds}, in a subsequent section. Here we only briefly mention two examples: Let $f, g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined by $f: (x, y) \mapsto x^2 + y^2 - 1$ and $g: (x, y) \mapsto xy$. Then $f^{-1}(0)$ is the 1-sphere, while $g^{-1}(0)$ is \emph{not} a manifold. The standard criterion that distinguishes these examples relies on derivatives: While $(Df)$ never vanishes where $f$ vanishes, $(Dg)$ and $g$ have common zeros – which are potentially troublesome points.

- There is a very rich world of complex manifolds – but we will not have the opportunity to explore this in any depth in this course. Within the frame of this section – coordinate charts for topological manifolds – complex manifolds do not offer any new features. But in the framework of differentiable manifolds, the much richer structure of complex differentiability opens completely new worlds, far beyond our course . . .
2.4 Differentiation: Notation and review

This subsection fixes some notation for partial derivatives of maps between Euclidean spaces, and contrasts this with a coordinate-free description of differentiation.

Let \(e_i = (0, 0, \ldots, 0, 1, 0 \ldots)^\top \in \mathbb{R}^n\) denote the standard \(i\)-th basis vector. For a function \(f: U \to \mathbb{R}\) defined on an open subset \(U \subseteq \mathbb{R}^n\) and \(a \in U\) the \(i\)-the partial derivative of \(f\) at \(a\) is defined (and denoted) by

\[
(D_i f)(a) = \lim_{h \to 0} \frac{1}{h} (f(a + he_i) - f(a))
\]

provided the limit exists. To denote different degrees of regularity we use the following notation:

- \(f \in C^0(U)\) if \(f\) is continuous on \(U\) (i.e., \(f\) is continuous at all \(p \in U\))
- \(f \in C^r(U)\) if all partial derivatives of up to order \(r\) of \(f\) exist and are continuous on \(U\),
- \(f \in C^\infty(U)\) if all partial derivatives of all orders of \(f\) exist and are continuous on \(U\),
- \(f \in C^r(\mathbb{R}^n)\) if \(f\) is real analytic on \(U\) (\(f \in C^\infty(U)\) and \(f\) agrees (locally) with its Taylor series).

For a function \(f: A \to \mathbb{R}\) defined on a set \(A \subseteq \mathbb{R}^m\) we say \(f \in C^\alpha(A)\) if there exist an open set \(O \subseteq \mathbb{R}^m\) such that \(A \subseteq O\) and an extension \(\bar{f}\) of \(f\) to \(O\) (i.e. \(\bar{f}|_A = f\)), and \(f \in C^\alpha(O)\).

If \(f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m\), then \(f \in C^\alpha(A, \mathbb{R}^m)\) if each coordinate function \(f^k = x^k \circ f \in C^\alpha(A)\).

If \(f \in C^1(\mathbb{R}^n, \mathbb{R}^m)\) then the Jacobian matrix is

\[
(Df) = (D_j f^i)_{i=1, \ldots, m \atop j=1, \ldots, n} = \begin{pmatrix} D_1 f^1 & \cdots & D_n f^1 \\ \vdots & \ddots & \vdots \\ D_1 f^m & \cdots & D_n f^m \end{pmatrix}
\]  

(13)

For convenience we identify the space \(M_{m,n}(\mathbb{R})\) of real \(m \times n\) matrices with \(\mathbb{R}^{mn}\). Thus, if \(f \in C^r(\mathbb{R}^n, \mathbb{R}^m)\), then \((Df) \in C^{r-1}(\mathbb{R}^n, \mathbb{R}^{mn})\), and \((D^k f) \in C^{r-k}(\mathbb{R}^n, \mathbb{R}^{mn})\).

The chain rule asserts that if \(U \subseteq \mathbb{R}^m\) and \(V \subseteq \mathbb{R}^n\) are open sets, \(g \in C^r(U, \mathbb{R}^n)\) and \(f \in C^r(V, \mathbb{R}^p)\) with \(r \geq 1\), \(g(U) \subseteq V\) and \(a \in U\), then \(f \circ g\) is differentiable at \(a\) and \(D(f \circ g)(a) = (Df)(g(a)) \cdot (Dg)(a)\) (matrix-multiplication).

Theorem 2.1 (Implicit function theorem) Suppose \(U \subseteq \mathbb{R}^{m+n}\) is open, \((a, b) \in U\) and \(f \in C^r(U, \mathbb{R}^m)\) with \(r \geq 1\), and \(f(a, b) = 0\). If the matrix of partial derivatives \((D_{a+j} f^i(a, b))_{i,j=1, \ldots, m}\) is nonsingular then there exist open sets \(V \subseteq \mathbb{R}^n\) and \(W \subseteq \mathbb{R}^m\) with \(a \in V\) and \(b \in W\) and a unique function \(g: V \to W\) such that \(f(x, g(x)) = 0\) for all \(x \in V\). Moreover, \(g \in C^r(V, W)\).

Differentiability (as opposed to mere existence of partial derivatives) may be nicely described in a coordinate-free way. Consider finite dimensional normed linear spaces \(V, W\) and let \(U \subseteq V\) be open. ((A norm is a map \(\| \cdot \|: V \to \mathbb{R}\) such that \(\|v\| \geq 0\) for all \(v \in V\), \(\|v\| = 0\) if and only of \(v = 0\), \(\|\lambda v\| = |\lambda| \|v\|\) for all \(v \in V\) and \(\lambda \in \mathbb{R}\), and \(\|v + w\| \leq \|v\| + \|w\|\) for all \(v, w \in V\).) As finite dimensional linear spaces \(V\) and \(W\) are isomorphic to some Euclidean spaces \(\mathbb{R}^n\) and \(\mathbb{R}^m\), but here the objective is to not fix any bases.

Definition 2.14 A map \(f: U \to W\) is differentiable at \(a \in U\) if there exists a linear map \(L = L_a f \in \text{Hom}(V, W)\) such that \(f(a + h) = f(a) + L(h) + o(||h||)\). (This means that there exists a map \(\eta: U \to W\) (depending on \(a\) and \(f\)) such that \(f(a + h) = f(a) + L(h) + \|h\| \eta(h)\) and \(\|\eta(h)\| \to 0\) at \(\|h\| \to 0\).)
Exercise 2.21 Show that if \( f \) is differentiable at \( a \) then the linear map of the preceding definition is uniquely determined. (Suppose there are two such linear maps. Show that their difference satisfies \( L(h) - M(h) = o(\|h\|) \). Therefore it is justified to talk about the derivative of \( f \) at \( a \) – and we will use the notation \( f'(a) \).

A map \( f \) is called differentiable on an open set \( U \) if \( f \) is differentiable at all \( a \in U \).

Note that if \( f: U \subseteq V \mapsto W \) then \( f': U \mapsto \text{Hom}(V, W) \). But the space \( \text{Hom}(V, W) \) of linear maps from \( V \) to \( W \) is itself a linear space. Hence one may naturally define higher order derivatives as linear maps \( f''': U \mapsto \text{Hom}(V, \text{Hom}(V, W)) \) and inductively \( f^{(k)}: U \mapsto \text{Hom}(\bigotimes_{i=1}^k V, W) \). (For comparison, if \( f: \mathbb{R}^n \mapsto \mathbb{R}^m \) then \( (Df) \) is an \( (m \times n) \) matrix, and, naively, \( (D^2 f) \) is some sort of \( (m \times n \times n) \) object which takes three inputs, a point where it is evaluated and two vectors . . . )

We summarize a few basic properties

(i) If \( f \) is differentiable then \( f \) is continuous.

(ii) If \( f \) and \( g \) are differentiable and \( \lambda \in \mathbb{R} \) then \( (f + g)' = f' + g' \) and \( (\lambda f)' = \lambda f' \).

(iii) If \( f \) is constant then \( f' \equiv 0 \).

(iv) If \( L \in \text{Hom}(V, W) \), \( b \in W \) and \( f = L + b \) (i.e. \( f: v \mapsto L(v) + b \)) then \( f' = L \).

(v) (Chain-rule). Let \( V_1, V_2, V_3 \) be normed linear spaces. Suppose \( U_1 \subseteq V_1 \) and \( U_2 \subseteq V_2 \) are open, \( g: U_1 \mapsto V_2 \), \( f: U_2 \mapsto V_3 \), and \( g(U_1) \subseteq U_2 \). If \( f \) and \( g \) are differentiable, then so is \( f \circ g \) and \( (f \circ g)' = (f' \circ g) \cdot g' \).

This notation is to be interpreted as follows: For \( p \in U_1 \) and \( v \in V_1 \), \( g'(p) \in \text{Hom}(V_1, V_2) \) and hence \( g'(p)(v) \in V_2 \). Similarly, \( f'(g(p)) \in \text{Hom}(V_2, V_3) \) and hence \( f'(g(p))(g'(v)) \in V_3 \). This matches with \( (f \circ g)'(p) \in \text{Hom}(V_1, V_3) \) and hence \( (f \circ g)'(p)(v) \in V_3 \).

2.5 Differentiable structures

In general, manifolds do not have a linear (or even only an additive structure). Thus expressions reminiscent of \( f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} \) are meaningless for maps \( f: M \mapsto N \) between manifolds (unless one considers points in \( M \) as distributions on the algebra of smooth functions . . .)

A natural way to define a notion of differentiability on manifolds is to utilize coordinate charts \((u, U)\) on \( M \) and \((v, V)\) on \( N \) to relate \( f: U \mapsto V \) to the map \( v \circ f \circ u^{-1} \) between Euclidean spaces. The main concern is to ensure that any so-defined notion of differentiation on a manifold does not depend on the particular choice of coordinates. This leads naturally to the concept of differentiable structures.

Definition 2.15 Two charts \((u_1, U_1)\) and \((u_2, U_2)\) on a manifold \( M \) are \( C^r \)-related \((r = 1, 2, \ldots, \infty, \omega)\) if the maps \( u_2 \circ u_1^{-1} \) and \( u_1 \circ u_2^{-1} \) are \( C^r \)-maps (as maps between Euclidean spaces) on their respective domains \( u_1(U_1 \cap U_2) \) and \( u_2(U_1 \cap U_2) \).

Definition 2.16 A \( C^r \)-differentiable structure on a manifold \( M \) is a maximal atlas \( D \) (i.e. a maximal collection of coordinate charts that covers \( M \)) such that any two charts \((u_1, U_1), (u_2, U_2)\) \( \in D \) are \( C^r \)-related. A manifold \( M \) together with a \( C^r \)-differentiable structure \( D \) is called a \( C^r \)-manifold, or simply, a differentiable manifold (if \( r \) is understood, usually \( r = \infty \)).
Definition 2.17 Two $C^r$-manifolds $(M, \mathcal{D})$ and $(N, \mathcal{D}')$ are called diffeomorphic if there exists a bijection $\Phi: M \mapsto N$ such that $(v, V) \in \mathcal{D}'$ if and only if $(v \circ \Phi, \Phi^{-1}(V)) \in \mathcal{D}$. The map $\Phi$ is called a diffeomorphism.

It is easy to see that every diffeomorphism must be continuous. Since the inverse $\Phi^{-1}$ is automatically a diffeomorphism, $\Phi$ is automatically a homeomorphism. However, manifolds may be homeomorphic without being diffeomorphic, see below.

Proposition 2.2 Every $C^r$-atlas is contained in a unique $C^r$-differentiable structure.

Proof. Let $\mathcal{U} = \{(u_\alpha, U_\alpha) : \alpha \in A\}$ be a $C^r$-atlas for a manifold $M$ — i.e. any two charts $(u_\alpha, U_\alpha), (u_\beta, U_\beta) \in \mathcal{U}$ are $C^r$-related and $M \subseteq \bigcup_{\alpha \in A} U_\alpha$. Define $\mathcal{D}$ to be the collection of all coordinate charts $(v_\alpha, V_\alpha)$ on $M$ which (each) are $C^r$-related to every $(u_\alpha, U_\alpha) \in \mathcal{U}$.

Maximality of $\mathcal{D}$ is clear. To verify that any two charts $(v_\alpha, V_\alpha), (v_\beta, V_\beta) \in \mathcal{D}$ are $C^r$-related it suffices to show that $v_\beta \circ v_\alpha^{-1} : v_\alpha(V_\alpha \cap V_\beta) \mapsto \mathbb{R}^m$ is locally $C^r$. Thus suppose that $x \in v_\alpha(V_\alpha \cap V_\beta) \subseteq \mathbb{R}^m$ and let $p = v_\alpha^{-1}(x) \in M$. Since $\mathcal{U}$ is an atlas of $M$, it contains a chart $(u, U) \in \mathcal{U}$ about $p$. On the set $v_\alpha(U \cap V_\alpha \cap V_\beta)$ we may write

$$v_\beta \circ v_\alpha^{-1} = (v_\beta \circ u^{-1}) \circ (u \circ v_\alpha^{-1})$$

as a composition of $C^r$-maps — and hence $v_\beta \circ v_\alpha^{-1}$ is locally $C^r$ on $v_\alpha(U \cap V_\alpha \cap V_\beta)$. Since the latter set is open in $\mathbb{R}^m$, $v_\beta \circ v_\alpha^{-1}$ is $C^r$.

Regarding uniqueness, suppose $\mathcal{D}'$ is any $C^r$ differentiable structure containing $\mathcal{U}$. Then by definition every $(v, V) \in \mathcal{D}'$ is $C^r$-related to every $(u, U) \in \mathcal{U}$, and consequently $(v, V) \in \mathcal{D}$, i.e. $\mathcal{D}' \subseteq \mathcal{D}$. Since a differentiable structure is maximal by definition, also $\mathcal{D} \subseteq \mathcal{D}'$, i.e. $\mathcal{D} = \mathcal{D}'$. $\blacksquare$

Definition 2.18 Let $(M, \mathcal{D})$ and $(N, \mathcal{D}')$ be differentiable manifolds of class $C^r$ and $C^s$, respectively, and $k \leq \min\{r, s\}$. A map $\Phi: M \mapsto N$ is called differentiable of class $C^k$ if for any charts $(u, U) \in \mathcal{D}$ and $(v, V) \in \mathcal{D}'$ the map $v \circ \Phi \circ u^{-1}$ is of class $C^k$ on its domain.

- If $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$ (with the usual differentiable structures) then $\Phi: M \mapsto N$ is differentiable in the usual sense.
- A map $\Phi: M \mapsto \mathbb{R}^n$ is differentiable if and only if each coordinate function $\Phi^k = x^k \circ \Phi$ is differentiable.
- Coordinate maps $u$ are diffeomorphisms from $U$ to $u(U)$.
- A map $\Phi: M \mapsto \mathbb{R}^n$ is a diffeomorphism if and only if it is bijective, differentiable, and its inverse is differentiable.

Proposition 2.3 Every differentiable map $\Phi: M \mapsto N$ between manifolds is continuous.

Proof. Let $W \subseteq N$ be open. We will show that $\Phi^{-1}(W) \subseteq M$ is open. Let $\mathcal{U} = \{(u_\alpha, U_\alpha) : \alpha \in A\}$ and $\mathcal{V} = \{(v_\beta, V_\beta) : \beta \in B\}$ be atlases for $M$ and $N$, respectively. Since $W = \bigcup_{\beta \in B} (V_\beta \cap W)$ we have $\Phi^{-1}(W) = \bigcup_{\beta \in B} \Phi^{-1}(V_\beta \cap W)$. Using that $u_\alpha$ is one-to-one for each $\alpha \in A$

$$u_\alpha(U_\alpha \cap \Phi^{-1}(W)) = u_\alpha(U_\alpha \cap \bigcup_{\beta \in B} \Phi^{-1}(V_\beta \cap W)) = \bigcup_{\beta \in B} u_\alpha(U_\alpha) \cap (u_\alpha \circ \Phi^{-1} \circ v_\beta^{-1}) \circ v_\beta(V_\beta \cap W)$$

(14)
Since \( v_\beta(V_\beta \cap W) \subseteq \mathbb{R}^n \) is open for each \( \beta \), and \((v_\beta \circ \Phi \circ u_\alpha^{-1})\) is continuous for each \( \alpha, \beta \), each of the sets \((u_\alpha \circ \Phi^{-1} \circ v_\beta^{-1}) \circ v_\beta(V_\beta \cap W)\) is open. Consequently, each \( u_\alpha(U_\alpha \cap \Phi^{-1}(W)) \subseteq \mathbb{R}^m \) is open, and thus each \( U_\alpha \cap \Phi^{-1}(W) \subseteq M \) is open. Finally, \( \Phi^{-1}(W) = (\bigcup U_\alpha \cap \Phi^{-1}(W) = = \bigcup U_\alpha (U_\alpha \cap \Phi^{-1}(W)) \) is a union of open sets, and thus is open. \[ \square \]

**Exercise 2.22** Show that the identity map \( \text{id}_M : M \rightarrow M \) on a \( C^r \)-manifold \( M \) is a \( C^r \)-map. Suppose that \( M, N, \) and \( P \) are \( C^r \)-manifolds with \( A \subseteq M \) and \( B \subseteq N \). Show that if \( g \in C^r(A, N) \) and \( f \in C^r(B, P) \) then \( f \circ g \in C^r(A \cap g^{-1}(B), P) \).

**Exercise 2.23** Consider \( M = \mathbb{R} \) with charts \((u_k, U)\), \( k = 1, 2, 3 \) where \( U = \mathbb{R} \) and \( u_1(x) = x^3 \), \( u_2(x) = x \), and \( u_3(x) = x^{1/3} \). In each case provide an example of another chart \((v_k, V_k)\) that is contained in the unique \( C^1 \)-differentiable structure \( D_k \) on \( M \) that contains \((u_k, U_k)\).

Explain why \( D_k \) are (pairwise) different, give examples of charts contained in their intersections (or explain why the intersections are empty). Demonstrate that \((M, D_k)\) are diffeomorphic.

**Exercise 2.24** (Continuation of exercise 2.23). Consider the maps \( \Phi_{i,j,k} : (M, D_i) \mapsto (M, D_j) \) defined by \( \Phi_{i,j,k}(x) = x^k \) for \( k = \frac{1}{3}, \frac{1}{2}, 1, 3, 9 \). Which of these maps are differentiable? Which maps are diffeomorphisms?

Note that every differentiable structure \( D \) of class \( C^r \), \( r \geq 1 \) may be regarded as an atlas of class \( C^{r-1} \), and hence is contained in a unique differentiable structure \( D' \) of class \( C^{r-1} \). Consequently the class of a manifold can be lowered at will, by adding new charts to an atlas.

More important is that every \( C^r \)-differentiable structure with \( r \geq 1 \) contains a \( C^\infty \) differentiable structure (see also below). Consequently one routinely restricts one’s attention to \( C^\infty \)-manifolds.

**Blanket hypothesis.** Unless otherwise stated, all manifolds and maps considered henceforth are assumed to be \( C^\infty \) manifolds and \( C^\infty \)-maps, respectively.

The following remarks are taken from my class-notes from UC Boulder in 1983 – they have not independently been verified (nor updated) ...

A (very hard) theorem by Whitney asserts that every \( C^1 \) differentiable structure contains a \( C^\omega \) structure. On the other hand, Kervaire has given an example of a 10 dimensional \( C^0 \) manifold which admits no differentiable structure (hard!).

The spheres \( S^n \) have unique differentiable structures for \( n \leq 6 \), but Milnor showed in 1958 that \( S^7 \) admits 28 non-diffeomorphic differentiable structures! However, for any \( n \) there is only a finite number of “diffeo-classes” on \( S^n \).

For \( n \neq 4 \) the Euclidean space \( \mathbb{R}^n \) has a unique differentiable structure, but \( \mathbb{R}^4 \) has at least 3 non-diffeomorphic differentiable structures (1982).
2.6 Partitions of unity

It is very common that one can easily construct objects locally, e.g. working on coordinate charts. Partitions of unity are a versatile tool to patch together such objects into a globally defined one. From a different point of view, partitions of unity demonstrate that there are plenty of $C^\infty$ functions on a differentiable manifold (as opposed to comparatively few $C^\infty$-functions).

We begin with some fundamental constructions in Euclidean spaces.

Lemma 2.4 Let $m \geq 1$, $0 < a < b$ and $p \in \mathbb{R}^m$. Then there exists a map $k \in C^\infty(\mathbb{R}^m, \mathbb{R})$ such that $k(x) = 0$ for $\|x-p\| \geq b$, $k(x) = 1$ for $\|x-p\| \leq a$, and $0 < k(x) \leq 1$ for $\|x-p\| \leq b$.

**Proof.** Let $f(t) = \exp(-1/t)$ for $t > 0$ and $f(t) = 0$ else. Then $f \geq 0$ and $f \in C^\infty(\mathbb{R})$. Next define

$$g(t) = \frac{f(t)}{f(t) + f(b-t)} \text{ if } t > 0$$

and $g(t) = 0$ for $t \leq 0$. Then $g(t) = 1$ for all $t \geq b$ and $g'(t) > 0$ for $0 < t < b$. Define

$$h(p) = g\left(\frac{b(b+t)}{b-a}\right) \cdot g\left(\frac{b(b-t)}{b-a}\right).$$

and finally set $k(x) = h(\|x-p\|)$. ■

One may replace the balls $B_p(r)$ in the lemma by cubes $C_p(r) = \{x \in \mathbb{R}^m : |x_i - p_i| \leq r\}$ by taking $k(x) = h(x_1 - p_1) \ldots h(x_m - p_m)$.  

**Exercise 2.25** Prove that $f \in C^\infty(\mathbb{R})$ as claimed in the preceding proof. (Use induction.)

**Proposition 2.5** Let $M^m$ be a $C^\infty$-manifold, $V \subseteq M$ open, and $K \subseteq V$ compact. Then there exists a function $\phi \in C^\infty(M, [0,1])$, $\phi|_K \equiv 1$ and $\phi|_W \equiv 0$ for some open set $W \supseteq M - V$.

**Proof.** Use the compactness of $K$ to select charts $(u_i, U_i)$, $i = 1, \ldots, N_1$ such that $K \subseteq \bigcup_{i=1}^{N_1} U_i$. Since $u_i(U_i \cap V) \subseteq \mathbb{R}^m$ is open (and w.l.o.g. nonempty), there exist $r_{y,i} > 0$ for every $y \in u_i(U_i \cap V)$ such $C_{y_i}(r_{y,i}) \subseteq u_i(U_i \cap V)$. The collection $\{u_i^{-1}(C_{y_i}(r_{y,i})); i \leq N_1, u_i^{-1}(y) \in K\}$ is an open cover of $K$.

Choose a finite subcover $\{u_i^{-1}(C_{y_{i,j}}(r_{i,j})); i \leq N_1, j \leq N_2(i)\}$, (writing $r_{i,j}$ for $r_{y_{i,j},i}$). By the preceding proposition there exist functions $h_{i,j}: \mathbb{R}^m \rightarrow [0,1]$ such that $h_{i,j}(x) = 1$ for all $x \in C_{y_{i,j}}(\frac{1}{4}r_{i,j})$ and $h_{i,j}(x) = 0$ for all $x \notin C_{y_{i,j}}(\frac{3}{4}r_{i,j})$. Define $\phi_{i,j} = h_{i,j} \circ u_i$ on $U_i$ and extend to $M$ by setting $\phi_{i,j}(q) = 0$ for $q \notin U_i$. Combine these functions into

$$\phi(x) = 1 - \prod_{i,j} (1 - \phi_{i,j}(x))$$

and set

$$W = M \setminus \bigcup_{i,j} u_i^{-1}\left(C_{y_{i,j}}(\frac{3}{4}r_{i,j})\right) \supseteq M \setminus V$$

Note that $\phi|_K \equiv 1$ since $K \subseteq \bigcup_{i,j} u_i^{-1}(C_{y_{i,j}}(\frac{1}{4}r_{i,j}))$. Clearly $\phi|_W \equiv 0$. ■

The objective is to use these bump-functions to patch together local results. It is a natural to require that at any fixed point only a finite number of the local results are needed, or may be
selected. This is a place where our assumptions come into play that a manifold’s topology be reasonably nice. For example, every metric (metrizable) space is paracompact (Stone’s theorem), and hence normal. The following shrinking lemma is a direct consequence of these properties. The construction in its proof is a good exercise to practice working with paracompactness and normality, and a good check for understanding. The lemma itself plays a fundamental role in the desired partitions of unity which are used to patch together the local results.

Before proceeding with the shrinking lemma, we provide a few optional side-remarks.

Compactness may be characterized in the following way, which may seem unusual, but which lends itself a natural generalization: “A space \( K \) is compact if every open cover of \( X \) has a finite open refinement that covers \( X \).” From here it is only a small step to paracompactness, which weakens “finite” to “locally finite” (and traditionally explicitly requires that the space is Hausdorff). According to Munkres (Topology, a first course): “The concept of paracompactness is one of the most useful generalizations of compactness that has been discovered in recent years. Particularly is it useful for applications in algebraic topology, differential geometry, . . . ” In point set topology its close connection with metrizability is utilized.

The following simple example illustrates how the definition works. Consider the real line, \( \mathbb{R} \), which lends itself a natural generalization:

\[
\text{Compactness may be characterized in the following way, which may seem unusual, but which lends itself a natural generalization: “A space } K \text{ is compact if every open cover of } X \text{ has a finite open refinement that covers } X. \text{.” From here it is only a small step to paracompactness, which weakens “finite” to “locally finite” (and traditionally explicitly requires that the space is Hausdorff). According to Munkres (Topology, a first course): “The concept of paracompactness is one of the most useful generalizations of compactness that has been discovered in recent years. Particularly is it useful for applications in algebraic topology, differential geometry, . . . ” In point set topology its close connection with metrizability is utilized.}
\]

The following simple example illustrates how the definition works. Consider the real line, \( \mathbb{R} \), which lends itself a natural generalization:

\[
\text{Compactness may be characterized in the following way, which may seem unusual, but which lends itself a natural generalization: “A space } K \text{ is compact if every open cover of } X \text{ has a finite open refinement that covers } X. \text{.” From here it is only a small step to paracompactness, which weakens “finite” to “locally finite” (and traditionally explicitly requires that the space is Hausdorff). According to Munkres (Topology, a first course): “The concept of paracompactness is one of the most useful generalizations of compactness that has been discovered in recent years. Particularly is it useful for applications in algebraic topology, differential geometry, . . . ” In point set topology its close connection with metrizability is utilized.}
\]
Continue by inductively selecting open sets $V_k \subseteq M$ such that $F_k \subseteq V_k \subseteq \tilde{V}_k \subseteq U_k$ and defining $F_{k+1} = U_{k+1} \setminus \bigcup_{i \leq k} V_i \cup \bigcup_{i > k+1} U_i$. At each stage $\{V_i : i \leq k\} \cup \{U_i : i > k+1\}$ is an open cover of $M$. Therefore each $F_{i+1} \subseteq M$ is closed.

We verify that the collection $\{V_i : i \geq 1\}$ still covers $M$: Suppose $p \in M$. Since $\mathcal{U}$ is locally finite, there exists for a finite number $n(p)$ such that $p \notin U_i$ for all $i > n(p)$. Consequently, $p \notin V_i$ for all $i > n(p)$. In other words, $p \in \bigcup_{i=1}^{n(p)} V_i$. 

**Theorem 2.7 (Partition of unity)** Let $\mathcal{U} = \{U_{\alpha}\}_{\alpha \in A}$ be an open cover of a manifold $M$. Then there exist $C^\infty$-functions $\phi_\alpha : M \mapsto [0,1]$ such that the collection of closed sets $W_\alpha = \{p \in M : \phi(p) \neq 0\}$, $\alpha \in A$ is locally finite, $W_\alpha \subseteq U_\alpha$ for each $\alpha \in A$, and $\sum_{\alpha \in A} \phi_\alpha \equiv 1$.

The collection $\{\phi_\alpha\}_{\alpha \in A}$ is called a partition of unity subordinate to $\mathcal{U}$.

**Proof.** Using the paracompactness of $M$ we may assume that $\mathcal{U}$ is locally finite. Else, find a locally finite open refinement $\tilde{\mathcal{U}}$ that covers $M$, and relabel $\tilde{\mathcal{U}}$ to be $\mathcal{U}$.

First consider the special case that $\tilde{U}_\alpha \subseteq M$ is compact for each $\alpha \in A$. Applying the shrinking lemma, obtain an open cover $\mathcal{V} = \{V_{\alpha}\}_{\alpha \in A}$ such that $\tilde{V}_\alpha \subseteq U_\alpha$ for each $\alpha \in A$.

As a closed subset of the compact set $\tilde{U}_\alpha$ the set $V_\alpha$ is again compact, and proposition 2.6 applies. Thus there are maps $\psi_\alpha \in C^\infty(M)$ such that $\psi_\alpha|_{\tilde{V}_\alpha} \equiv 1$. Moreover, if $Z_\alpha = \{x \in M : \psi_\alpha(x) \neq 0\}$ then $\bar{Z}_\alpha \subseteq U_\alpha$. (Using the notation of the proposition 2.6, $Z_\alpha \subseteq W_\alpha$, and since each $W_\alpha$ is closed, it follows that $\bar{Z}_\alpha \subseteq W_\alpha$, and hence $\bar{Z}_\alpha \subseteq U_\alpha$.)

Every point $p \in M$ has an open neighbourhood $O \subseteq M$ that meets only finitely many $U_\alpha$. Consequently, all but a finite number of the functions $\psi_\alpha$ vanish identically on $O$, and the sum $\sum_{\alpha \in A} \psi_\alpha$ is well-defined on $O$, and hence on all of $M$. Moreover, since $\mathcal{V}$ is a cover for $M$, for every point $p$ in $M$ there exists some $\alpha \in A$ such that $p \in V_\alpha$ and thus $\psi_\alpha(p) > 0$. Define

$$
\phi_\alpha = \frac{\psi_\alpha}{\sum_{\beta \in A} \psi_\beta}.
$$

Clearly $0 \leq \phi_\alpha \leq 1$ for all $\alpha \in A$ and $\sum_{\alpha \in A} \phi_\alpha \equiv 1$. Moreover the support $\text{supp}(\phi_\alpha) = \{x \in M : \phi(x) \neq 0\}$ is contained in $U_\alpha$ since $\text{supp}(\phi_\alpha) \subseteq Z_\alpha \subseteq \bar{Z}_\alpha \subseteq U_\alpha$.

We now use this special case to prove a strengthened version of proposition 2.6, and then use that strengthened version to prove the existence of a partition of the unity in the general case.

Suppose $F \subseteq M$ is closed (not necessarily compact), $O \subseteq M$ open and $F \subseteq O$. For each $x \in F$ choose an open neighbourhood $V(x) \subseteq O$ such that $\overline{V(x)}$ is compact. For each $x \notin F$ choose an open neighbourhood $V(x)$ such that $F \cap \overline{V(x)} = \emptyset$ and such that $\overline{V(x)}$ is compact. (This uses the normality of $M$.) The open cover $\{V(x)\}_{x \in M}$ has a locally open refinement $\{Z(x)\}_{x \in M}$ that covers $M$. (Note that $\overline{Z(x)} = \emptyset$ may happen for many $x \in M$.)

Since the sets $\overline{Z(x)}$ are compact, the special case of the partition of the unity theorem applies. This means that there are functions $\phi_x \in C^\infty(M, [0,1])$ such that the collection $\{y \in M : \phi_x(y) > 0 : x \in M\}$ is locally finite, $\{y \in M : \phi_x(y) > 0\} \subseteq Z(x)$ and $\sum_{x \in M} \phi_x \equiv 1$.

Set $f = \sum_{x \in F} \phi_x$. Clearly $f \in C^\infty(F, [0,1])$. If $x \notin F$, then $\text{supp}(\phi_x) \subseteq \overline{Z(x)} \subseteq \overline{V(x)} \subseteq M \setminus F$. Consequently, $f|_{F} \equiv 1 - \sum_{x \notin F} \phi_x|_{F} \equiv 1 - 0 = 1$. 

In the next subsection we will use the partition of unity theorem to prove that every compact $C^\infty$ manifold may be embedded in some Euclidean space. (This is even true without the compactness assumption, but considerably harder to prove.)
2.7 Differentiating differentiable maps

We start this section by defining (partial) derivatives of maps on coordinate charts. This notion will enable us to define submersions, immersions, and embeddings. These in turn are useful tools to construct new differentiable manifolds from known ones. As examples we revisit hypersurfaces and some matrix submanifolds, and we show that every compact manifold can be imbedded into a Euclidean space.

Recall the definition of the \( i \)-th partial derivative of a function \( f: \mathbb{R}^n \to \mathbb{R} \) at a point \( a \in \mathbb{R}^n \):
\[
D_i f(a) = \lim_{h \to 0} \frac{1}{h} (f(a_1, \ldots, a_{i-1}, a_i + h, a_{i-1}, \ldots, a_n) - f(a)).
\]

**Definition 2.19** For a function \( f: M \to \mathbb{R} \) the \( i \)-th partial derivatives in a local coordinate chart \((u, U)\) at a point \( p \in U \) is defined as
\[
\frac{\partial f}{\partial u^i} \bigg|_p = D_i(f \circ u^{-1}) \bigg|_{u(p)}
\]

Note, if \( \gamma: (-\varepsilon, \varepsilon) \to U \subseteq M \) is a curve such that \( \gamma(0) = p \), \( u^i(\gamma(t)) = t \) and \( u^j(\gamma(t)) = u^j(p) \) for \( j \neq i \) then for \( f: U \to \mathbb{R} \)
\[
\frac{\partial f}{\partial u^i} \bigg|_p = D_i(f \circ u^{-1}) \bigg|_{u(\gamma(0))} = \lim_{h \to 0} \frac{1}{h} (f(\gamma(h)) - f(p))
\]

Note that this curve \( \gamma \) is adapted in a very special way to the specific local coordinate chart. The next section will generalize this setting along its way to construct the tangent bundle.

- If \( f = u^j \) then \( \frac{\partial u^i}{\partial u^j} = \delta_{i,j} \) (Kronecker delta, i.e., \( \delta_{i,j} = 1 \) if \( i = j \) and \( \delta_{i,j} = 0 \) else).
- If \( M = \mathbb{R}^n \) with the identity chart then \( \frac{\partial f}{\partial a^i} = D_i f \).

**Proposition 2.8 (Chain-rule)** Let \((u, U)\) and \((v, V)\) be local coordinate charts, \( p \in U \cap V \) and \( f: M \to \mathbb{R} \). Then
\[
\frac{\partial f}{\partial v^a} \bigg|_p = \sum_{j=1}^n \frac{\partial f}{\partial u^j} \bigg|_p \frac{\partial u^j}{\partial v^a} \bigg|_p
\]

**Proof.** Use the coordinate maps and the chain-rule in Euclidean spaces
\[
\frac{\partial f}{\partial v^a} \bigg|_p = D_i(f \circ v^{-1}) \bigg|_{v(p)} = D_i((f \circ u^{-1}) \circ (u \circ v^{-1})) \bigg|_{v(p)} = \sum_{j=1}^n D_j((f \circ u^{-1}) \circ (u \circ v^{-1})) \bigg|_{v(p)} \cdot D_i(u^j \circ v^{-1}) \bigg|_{v(p)} = \sum_{j=1}^n \frac{\partial f}{\partial u^j} \bigg|_p \frac{\partial u^j}{\partial v^a} \bigg|_p. \quad \blacksquare
\]

**Proposition 2.9 (Product-rule)** Suppose \((u, U)\) is a chart, \( p \in U \) and \( f, g: M \to \mathbb{R} \). Then
\[
\frac{\partial (fg)}{\partial v^a} \bigg|_p = \frac{\partial f}{\partial v^a} \bigg|_p \cdot g(p) + f(p) \cdot \frac{\partial g}{\partial v^a} \bigg|_p
\]

**Exercise 2.26** Prove the product rule for functions on a manifold as stated above.
Proposition 2.10 Let \((u,U)\) and \((v,V)\) be local coordinate charts, about \(p \in U \cap V\). Then the Jacobian matrix
\[
\left( \frac{\partial u^i}{\partial v^j} \right)_{p} \quad \text{is invertible.} \tag{25}
\]

**Proof.** The chain-rule gives \(\delta_{i,k} = \frac{\partial u^i}{\partial v^j} \frac{\partial v^j}{\partial \bar{u}^k} \big|_{p} = \sum_{j=1}^{n} \frac{\partial u^i}{\partial \bar{u}^j} \frac{\partial \bar{u}^j}{\partial v^k} \big|_{p} \), Hence the Jacobian matrices
\[
A = \left( \frac{\partial u^i}{\partial v^j} \right)_{p} \quad \text{and} \quad B = \left( \frac{\partial v^j}{\partial \bar{u}^k} \right)_{p} \quad \text{have the same rank} \tag{26}
\]

Proposition 2.11 Let \(\Phi: M \to N\) be a differentiable map between manifolds. If \((u,U)\) and \((\bar{u},\bar{U})\) are local coordinate charts about \(p \in M^m\), and \((v,V)\) and \((\bar{v},\bar{V})\) are local coordinate charts about \(\Phi(p) \in N^n\), then the Jacobian matrices (at \(p \in U \cap \bar{U}\))
\[
\left( \frac{\partial (v^i \circ \Phi)}{\partial \bar{u}^j} \right)_{p} \quad \text{and} \quad \left( \frac{\partial (\bar{v}^j \circ \Phi)}{\partial u^i} \right)_{p} \quad \text{have the same rank} \tag{27}
\]

**Proof.** Applying the chain-rule twice yields
\[
\frac{\partial (\bar{v}^j \circ \Phi)}{\partial \bar{u}^j} \big|_{p} = D_j (\bar{v}^j \circ \Phi \circ \bar{u}^{-1}) \big|_{\bar{u}(p)} = D_j \left( (\bar{v}^j \circ v^{-1}) \circ (v \circ \Phi \circ u^{-1}) \circ \left( u \circ \bar{u}^{-1} \right) \right) \big|_{\bar{u}(p)} = \sum_{k=1}^{m} \sum_{\ell=1}^{n} D_k \left( \bar{v}^j \circ v^{-1} \right) \big|_{v(p)} \cdot D_{\ell} \left( \bar{u}^k \circ \Phi \circ u^{-1} \right) \big|_{u(p)} \cdot D_j \left( \bar{u}^\ell \circ \bar{u}^{-1} \right) \big|_{\bar{u}(p)} \tag{28}
\]

Now use that the Jacobian matrices
\[
\left( \frac{\partial \bar{v}^j}{\partial \bar{u}^k} \right)_{\Phi(p)} \quad \text{and} \quad \left( \frac{\partial u^i}{\partial \bar{u}^j} \right)_{p} \quad \text{are invertible to conclude the argument.} \tag{29}
\]

Definition 2.20 Let \(\Phi: M \to N\) be a differentiable map between manifolds. The rank of \(\Phi\) at \(p \in M\) is defined as the rank of the Jacobian matrix
\[
\left( \frac{\partial (v^i \circ \Phi)}{\partial \bar{u}^j} \right)_{p} \quad \text{where} \quad (u,U) \quad \text{and} \quad (v,V) \quad \text{are any local coordinate charts about} \quad p \in M \quad \text{and} \quad \Phi(p) \in N, \quad \text{respectively.} \tag{30}
\]

This definition is justified – the rank is well-defined – by virtue of the preceding proposition: The rank of \(\Phi\) at \(p\) (written \(\text{rank}_p \Phi\)) is independent of the choice of local coordinates employed.
Definition 2.21 Suppose $\Phi: M \to N$ is a differentiable map between manifolds.

- $p \in M$ is called a regular point of $\Phi$ if $\operatorname{rank}_p\Phi = \dim N$.
- $p \in M$ is called a critical point of $\Phi$ if $\operatorname{rank}_p\Phi < \dim N$.
- $q \in N$ is called a regular value of $\Phi$ if every point $p \in \Phi^{-1}(q) \subseteq M$ is a regular point of $\Phi$.
- $q \in N$ is called a critical value of $\Phi$ if $q = \Phi(p)$ for some critical point $p$ of $\Phi$.

Note that every point $q \in N \setminus \Phi(M)$ is automatically a regular value of $\Phi$. Intuitively one expects that the set of critical values is a small subset of $N$. To make this precise it suffices to have a notion of sets of measure zero on the manifold. (This does not require a notion of a measure on the manifold!)

Definition 2.22 A subset $S \subseteq M^n$ of a manifold $M$ has measure zero if there exist a sequence of charts $(u_i, U_i), i \in \mathbb{Z}^+$ such that $S \subseteq \bigcup_{i=1}^{\infty} U_i$ and each $u_i(U_i) \subseteq \mathbb{R}^n$ has measure zero.

Exercise 2.27 Suppose $\Phi \in C^1(M, N)$ is a differentiable map between manifolds and $S \subseteq M$ has measure zero. Show that $\Phi(S) \subseteq N$ has measure zero.

Theorem 2.12 (Sard’s theorem, simple version) Suppose $M$ and $N$ are differentiable manifolds with $\dim M = \dim N$ and $M$ has at most countably many connected components. If $\Phi \in C^1(M, N)$ then the set of critical values of $\Phi$ has measure zero in $N$.

Note that with our definition of sets of measure zero this theorem is a direct consequence of the analogous statement for maps between Euclidean spaces. A proof of Sard’s theorem for such maps from $\mathbb{R}^n$ to $\mathbb{R}^n$ may be found in e.g. Spivak, Calculus on manifolds (p.72).

Definition 2.23 Suppose a differentiable map $\Phi: M^m \to N^n$ has constant rank $k$ at all $p \in M$.

- If $k = \dim N$ then $\Phi$ is called a submersion.
- If $k = \dim M$ then $\Phi$ is called an immersion and $\Phi(M) \subseteq N$ an immersed submanifold.
- If $\Phi$ is an immersion and $\Phi$ is also a homeomorphism onto its image $\Phi(M)$ (in the subspace topology inherited from $N$), then $\Phi$ is called an embedding, and $\Phi(M) \subseteq N$ is called an embedded submanifold, or simply a submanifold.

These definitions apply in particular to the case when the manifold $M$ is also a subset of $N$. In this case, the map $\Phi: M \to N$ is naturally taken as the inclusion map $\iota: M \hookrightarrow N$ defined by $\iota(p) = p$. Hence, the manifold $M$ is an immersed submanifold of $N$ if $\iota$ is an immersion, and $M$ is a submanifold of $N$ if $\iota$ is an imbedding. In the case that $M$ is a submanifold and is also a closed subset of $N$ one also calls $M$ a closed submanifold.

Note that an immersion is not required to be one-to-one. However, since it has maximal rank, the map is locally one-to-one. This means that every point $p \in M$ has a neighbourhood on which the map is one-to-one (i.e. the map is a topological immersion).
Examples (immersions and imbeddings):

- The map \( \Phi: \mathbb{R} \to \mathbb{R}, \Phi(x) = x^3 \) is \( C^\infty \) and is a homeomorphism, but not an immersion.
- The map \( g: \mathbb{R} \to \mathbb{R}^2, g(t) = (2\cos(t - \frac{\pi}{2}), \sin(2t - \pi)) \) is an immersion but not one-to-one.
- Define \( h: \mathbb{R} \to \mathbb{R} \) by \( h(t) = 2\tan^{-1}t \). Then \( (g \circ h): \mathbb{R} \to \mathbb{R}^2 \) is a one-to-one immersion, but not an imbedding: For every neighbourhood \( U \) of \( (g \circ h)(0) = (0, 0) \) in \( N = \mathbb{R}^2 \) there exists a \( a > 0 \) such that \( (-\infty, -a) \cup (a, \infty) \subseteq (g \circ h)^{-1}(U) \).
- A skew line on the torus is the image of a curve \( \gamma: \mathbb{R} \to \mathbb{R}^2/\mathbb{Z}^2 ; \gamma: t \mapsto (t, qt) \mod \mathbb{Z}^2 \) for an irrational number \( q \in \mathbb{R} \setminus \mathbb{Q} \). The irrationality assures that \( \gamma \) is one-to-one. But the image of \( \gamma \) is a dense subset on the torus, and hence it is not an embedding.

**Theorem 2.13** Suppose \( \Phi \in C^\infty (M^m, N^n) \) \( (r > 0) \) is a smooth map between manifolds and \( p \in M \). If \( \Phi \) has constant rank \( k \) on a neighbourhood of \( p \) then there exist local coordinate charts \((u, U)\) and \((v, V)\) about \( p \) and \( \Phi(p) \), respectively, such that for all \( x \in u(U) \subseteq \mathbb{R}^m \)

\[
v\left( (\Phi \circ u^{-1})(x^1, \ldots x^m) \right) = (x^1, \ldots x^k, 0, \ldots 0)
\]

**Proof.** Start with any charts \((\bar{u}, \bar{U})\) and \((\bar{v}, \bar{V})\) about \( p \) and \( \Phi(p) \), respectively. Without loss of generality assume that the first \((k \times k)\)-minor of the Jacobian has full rank, i.e.

\[
\det \left( \frac{\partial (\bar{v}^i \circ \Phi)}{\partial \bar{u}^j} \right)_{i=1,\ldots,k}^{j=1,\ldots,k} \neq 0 \tag{31}
\]

(else permute and relabel the components of \( \bar{u} \) and/or \( \bar{v} \)). This rank condition allows one to effectively use \((\bar{v}^i \circ \Phi), i = 1, \ldots k \) as coordinate functions in place of \( u^i \), \( i = 1, \ldots k \). More specifically define a new map \( \bar{u}: \bar{U} \to \mathbb{R}^m \) by \( \bar{u}^j = \bar{v}^j \circ \Phi \) if \( j \leq k \) and \( \bar{u}^j = \bar{u}^j \) else. To verify that this is indeed a legitimate local coordinate change calculate the Jacobian matrix of partial derivatives \( \frac{\partial \bar{u}^i}{\partial \bar{u}^j} \). This matrix has an upper block triangular structure. The first \((k \times k)\) minor agrees by construction with the one in (31), while the bottom right block is the \((m-k) \times (m-k)\) identity matrix. Consequently, the Jacobian has full rank at \( p \), and hence in some open neighbourhood of \( p \). By virtue of the inverse function theorem, there exists a neighbourhood \( U \subseteq \bar{U} \) of \( p \) such that the restriction \( u \) of \( \bar{u} \) to \( U \) is a diffeomorphism (onto its image), and hence \((u, U)\) is a chart about \( p \). Thus

\[
\bar{v} \left( (\Phi \circ u^{-1})(x^1, \ldots x^m) \right) = (x^1, \ldots x^k, \bar{v}^{k+1}(x), \ldots \bar{v}^n(x)) \tag{32}
\]

for suitable functions \( \bar{v}^{k+1}, \ldots, \bar{v}^n; \) \( u(U) \to \mathbb{R} \). The Jacobian matrix of partial derivatives \( (D_j(\bar{v}^i \circ \Phi \circ u^{-1})(x)), i = 1, \ldots, n, \) and \( j = 1, \ldots m \) has a lower block triangular structure. The top left \((k \times k)\) block is the identity, while the bottom right \((m-k) \times (n-k)\)-block consist of the partial derivatives \( D_j \bar{v}^i(x), i = k + 1, \ldots n \) and \( j = k + 1, \ldots m \). Since the matrix is assumed to have constant rank \( k \) in a neighbourhood of \( u(p) \in \mathbb{R}^m \), one concludes that the bottom right \((m-k) \times (n-k)\)-block is identically equal to zero. This means that the functions \( \bar{v}^i \) for \( i = k + 1, \ldots n \) do not depend on \( x^1, j = k + 1, \ldots m \). Consequently there are functions \( \bar{\psi}, i = k + 1, \ldots n, \) defined on a suitable subset of \( \mathbb{R}^k \) such that

\[
\bar{\psi}^i(x^1, \ldots, x^k, x^{k+1}, \ldots x^m) = \bar{v}^i(x^1, \ldots, x^k) \text{ for all } x = (x^1, \ldots x^m) \in u(U) \tag{33}
\]
Restrict the maps $\tilde{v}$ and $\tilde{\psi}$ to $V = \tilde{V} \cap \Phi(U)$, denoted $\tilde{v} = \tilde{v}_{|V}$ and $\psi = \tilde{\psi}_{|V}$. Finally define $v: V \to \mathbb{R}^n$ by $v^i = \tilde{v}^i$ if $i \leq k$ and $v^i = \tilde{v}^i - \psi^i \circ (v^1, \ldots, v^n)$ if $i > k$, i.e. for $y \in \tilde{v}(V \cap \Phi(U)) \subseteq \mathbb{R}^n$

$$
(v^i \circ \tilde{v}^{-1})(y_1, \ldots, y_n) = (y_1, \ldots, y^k, \psi^{k+1}(y_1, \ldots, y^k), \ldots, y^n - \psi^n(y_1, \ldots, y^k)) \tag{34}
$$

This assures that, as desired, for $x \in u(U \cap \Phi^{-1}(V))$

$$
(v \circ \Phi \circ u^{-1})(x^1, \ldots, x^m) = (v \circ \tilde{v}^{-1}) \circ (\tilde{v} \circ \Phi \circ u^{-1})(x^1, \ldots, x^m) = (v \circ \tilde{v}^{-1})(x^1, \ldots, x^k, \psi^{k+1}(x^1, \ldots, x^k), \ldots, \psi^n(x^1, \ldots, x^k)) = (x^1, \ldots, x^k, 0, \ldots, 0). \tag{35}
$$

Note that in general the construction in the proof works under the assumption that $\Phi \in C^r$ as long as $r > 0$. The strong version of the inverse function theorem yields local coordinates $u$ and $v$ of the same degree of smoothness as $\Phi$.

The special case of $k = \dim M$ is sometimes referred to as the “local immersion theorem”, while the special case of $k = \dim N$ is referred to as the “local submersion theorem”. In the first case the proof is a little shorter as it is immediately clear that $(v \circ \Phi)$ defines local coordinates about $p$. In the second case, the construction becomes shorter as there is no need for the functions $\tilde{\psi}$.

**Corollary 2.14** Suppose $\Phi \in C^\infty(M^m, N^n)$ is a smooth map between manifolds and $q \in N$. If $\Phi$ has constant rank $k$ on a neighbourhood of $\Phi^{-1}(q)$ then $\Phi^{-1}(q) \subseteq M$ is a closed submanifold of dimension $(m - k)$ (or it is empty). In particular, if $q \in N$ is a regular value of $\Phi$ then $\Phi^{-1}(q) \subseteq M$ is an $(m - n)$-dimensional submanifold on $M$.

**Proof.** Suppose $p \in \Phi^{-1}(q)$. Then there exist charts $(u, U)$ and $(v, V)$ about $p$ and $q$, respectively, such that w.l.o.g. $u(p) = 0 \in \mathbb{R}^m$, $v(q) = 0 \in \mathbb{R}^n$, and for $x \in u(U \cap \Phi^{-1}(V))$

$$
(v \circ \Phi \circ u^{-1})(x^1, \ldots, x^k, x^{k+1}, \ldots, x^m) = (x^1, \ldots, x^k) \tag{36}
$$

Let $W = U \cap \Phi^{-1}(q)$ and define $w: W \to \mathbb{R}^{m-k}$ by $w^j = u^{k+j}$ for $j = 1, \ldots, (m - k)$. Then $(w, W)$ is a chart for $\Phi^{-1}(M)$ about $p$. \hfill \blacksquare

This theorem is a special case of a more general result that applies to the inverse image $\Phi^{-1}(P) \subseteq M$ of a submanifold $P \subseteq N$ under a smooth map $\Phi: M \to N$. To illustrate that some additional hypotheses are needed, consider the example of $M = \mathbb{R}$, $N = \mathbb{R}^2$, $\Phi(x) = (x, 0)$ and $P = \{(x, f(x)): x \in \mathbb{R}\}$ is the graph of the $C^\infty$-function $f: \mathbb{R} \to \mathbb{R}$, defined by $f(x) = \exp(-\frac{1}{x})$ if $x > 0$ and $f(x) = 0$ else. Then $P \subseteq N$ is an (imbedded) 1-dimensional submanifold, $\Phi \in C^\infty(M, N)$ is a submersion (rank$_x\Phi = 1$ for all $x \in M$). However, $\Phi^{-1}(P) = (-\infty, 0) \subseteq M$ is not a submanifold (because no neighbourhood of $0 \in \Phi^{-1}(P) \subseteq M$ is homeomorphic to $\mathbb{R}$).

**Exercise 2.28** Prove the submersion theorem: If $\Phi \in C^\infty(M^m, N^n)$ is a submersion and $P \subseteq N$ is a submanifold of dimension $p$ then $\Phi^{-1}(P) \subseteq M$ is a submanifold of dimension $m - (n - p)$.

**Exercise 2.29** Explore where an attempt (using constructions similar to the ones used in the proofs of the preceding theorem and corollary) to prove the statement of the preceding exercise without the assumption that rank$\Phi = n$ breaks down (unless other assumptions are added).
Examples (Submanifolds from submersions):

- If \( f: \mathbb{R}^2 \rightarrow \mathbb{R} \), \( \Phi(x,y) = x^2 + y^2 \), then \( f^{-1}(1) = S^1 \subseteq \mathbb{R}^2 \) is an (imbedded) circle.
- If \( f: \mathbb{R}^{m+1} \rightarrow \mathbb{R} \), \( \Phi(x) = \|x\|^2 \), then \( f^{-1}(1) = S^m \subseteq \mathbb{R}^{m+1} \) is an (imbedded) m-sphere.
- Let \( \tilde{\Phi}: \mathbb{R}^4 \rightarrow \mathbb{R}^3 \) be defined by \( \tilde{\Phi}(a) = (2a_1a_3 + 2a_2a_4, 2a_2a_3 - 2a_1a_4, a_1^2 + a_2^2 - a_3^2 - a_4^2) \).

**Exercise 2.30** (Hopf map)
- Calculate the rank of \( \tilde{\Phi} \) at any \( a \in \mathbb{R}^4 \).
- Verify that \( \tilde{\Phi}(S^3) \subseteq S^2 \).
- Let \( \Phi \) be the restriction of \( \tilde{\Phi} \) to \( S^3 \subseteq \mathbb{R}^4 \), and considered as a map into \( S^2 \subseteq \mathbb{R}^3 \).
- Calculate the rank of \( \tilde{\Phi} \) at \( p \in S^3 \) using the definition in terms of coordinate charts. (Use stereographic projections to avoid square-roots!)
- For \( q \in S^2 \) describe the preimage \( \Phi^{-1}(q) \subseteq S^3 \). Is it a submanifold? If so, what is its dimension? Which manifold is it?

**Exercise 2.31** Define \( \Phi: M_{n \times n}(\mathbb{R}) \rightarrow M_{n \times n}(\mathbb{R}) \) by \( \Phi(A) = A^\top A \). Show that the set \( \Phi^{-1}(I_{n \times n}) \) of orthogonal matrices is a closed submanifold of \( M_{n \times n}(\mathbb{R}) \). What is its dimension?

- If \( \Phi = \det: M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R} \) is the determinant function, then the special linear group \( \text{SL}(n, \mathbb{R}) \) is the preimage \( \det^{-1}(1) \). The following calculations show that \( \text{SL}(n, \mathbb{R}) \) is a closed \((n^2 - 1)\)-dimensional submanifold of \( M_{n \times n}(\mathbb{R}) \).

Calculations for the example of the special linear group.

It suffices to show that \( \det \) has rank one at every \( A \in \det^{-1}(1) \subseteq M_{n \times n}(\mathbb{R}) \). The following calculations actually provide a little more, including a formula for the derivative of \( \det \) which is of interest in the next chapter. The calculation of the derivative in \( M_{n \times n}(\mathbb{R}) \) is basically the same as in \( \mathbb{R}^{n^2} \). Since we need inverses, we consider the restriction of \( \det \) to the the general linear group \( \text{GL}(n, \mathbb{R}) = \{ A \in M_{n \times n}(\mathbb{R}) : \text{det} A \neq 0 \} \) of invertible matrices. To make the connection with the next chapter, recall that \( \text{GL}(n, \mathbb{R}) \) is an open submanifold of \( M_{n \times n}(\mathbb{R}) \) – and thus it is appropriate to consider directional derivatives of at points \( A \in \text{GL}(n, \mathbb{R}) \) in all possible directions \( B \in M_{n \times n}(\mathbb{R}) \):

\[
f'(A)(B) = \lim_{h \to 0} \frac{1}{h} (f(A + hB) - f(A)) = \lim_{h \to 0} \frac{1}{h} (\det(A + hB) - \det A) = \lim_{h \to 0} \frac{1}{h} \cdot \det A \cdot \left( h^n \cdot \det \left( \frac{1}{h} I + A^{-1}B \right) - 1 \right) = \det A \lim_{h \to 0} \frac{1}{h} \left( h^n \cdot \left( \frac{1}{h^n} - \frac{1}{h^{n-1}} \cdot \text{tr}(A^{-1}B) + \ldots + (-)^n \det(A^{-1}B) \right) - 1 \right) = - \det A \cdot \text{tr}(A^{-1}B) \]

Choosing \( B = A \) shows \( f'(A)(A) = -n \det A \neq 0 \) and hence \( f'(A) \) has full rank equal to one, the dimension of the range. Thus \( \text{SL}(n, \mathbb{R}) \) is a an \((n^2 - 1)\)-dimensional closed submanifold.
As a final note in this chapter we combine imbeddings with a variation of partitions of the unity.

**Theorem 2.15** Any compact smooth manifold $M$ can be imbedded in some Euclidean space $\mathbb{R}^N$.

This is a very weak version – with a very simple proof – of much stronger results. Indeed, it is not too hard to actually show that it is always possible to take $N = 2n + 1$ (Spivak I, ex. 3.33). Much harder is Whitney’s theorem which asserts that actually $N = 2n$ is possible, even without the compactness assumption. A similar result for isometric (i.e. distance-preserving) imbeddings of Riemannian manifolds is known as Nash’s theorem.

**Proof.** Since $M^m$ is compact there exists a finite atlas $\mathcal{U} = \{(u_i, U_i): i = 1 \ldots n\}$ for $M$. Use the shrinking lemma to obtain an open cover $\mathcal{V}$ of $M$ by open sets $V_i \subseteq M$ whose closures are contained in $U_i$, i.e. $\overline{V}_i \subseteq U_i$, $i = 1, \ldots n$. Using normality, there are smooth functions $\phi_i: M \to [0, 1]$, $i = 1, \ldots n$ such that $\phi_i|_{\overline{V}_i} \equiv 1$ and such that each has support $\text{supp}(\phi_i) = \{p \in M: \phi_i(p) \neq 0\} \subseteq U_i$. Extend each map $u_i$ to all of $M$ by setting $u_i(p) = 0$ if $p \notin U_i$. Let $N = nm + n$ and define $\Phi: M \to \mathbb{R}^N$ by

$$\Phi = (\phi_1 \cdot u_1, \ldots, \phi_n \cdot u_n, \phi_1, \ldots, \phi_n)$$

To see that $\Phi$ is one-to-one, suppose $p, q \in M$ are such that $\Phi(p) = \Phi(q)$. Since $\mathcal{V}$ covers $M$ there exists $i_0$ such that $\phi_{i_0}(p) = 0$. Since $\Phi(p) = \Phi(q)$ this implies that $\phi_{i_0}(q) = \phi_{i_0}(p) = 0$, too. Consequently both $p \in U_{i_0}$ and $q \in U_{i_0}$. Again use $\Phi(p) = \Phi(q)$ to conclude that $\phi_{i_0}(q)u_{i_0}(q) = \phi_{i_0}(p)u_{i_0}(p)$, hence $u_{i_0}(q) = u_{i_0}(p)$, i.e. $p = q$.

Finally verify that $\Phi$ is an immersion: Let $p \in M$ and choose $i_0$, as above, that $p \in V_{i_0}$, i.e. $\phi_{i_0} \equiv 1$ in some neighbourhood of $p$. Consequently $\Phi^{m-(i_0-1)+j} \equiv u_{i_0}^j$, $j = 1, \ldots m$ in a neighbourhood of $p$. This guarantees that the $(N \times m)$-Jacobian $D(\Phi \circ u_{i_0}^{-1})$ contains an $(m \times m)$-identity block, i.e. $\Phi$ has full rank in a neighbourhood of $p$. □
3 The tangent bundle

3.1 Introduction

This is a good time to reflect why we want a notion of tangent spaces and tangent maps in the first place. Said differently, what do we expect this notion to deliver? What properties should tangent vectors and tangent spaces have? What are the tangent spaces to the line \( \mathbb{R} \) and the plane \( \mathbb{R}^2 \) — two of the most simple manifolds?

We want to use our experience with tangent lines to curves and tangent planes to surfaces in two- and three dimensional Euclidean spaces as guidance. However, in general we do not want our notion of tangent objects to depend on, or be constrained by imbeddings of the manifold into some Euclidean space. Thus without any surrounding space available, the pictorial arrows become untenable. Before reading on, you should close the notes and brainstorm some ideas . . .

Some ideas which come to mind are:

- Tangent vectors should be vectors, i.e. be members of a linear space that provides for addition and scalar multiplication.
- The dimension of the linear tangent space(s) should equal the dimension of the manifold.
- Tangent objects should provide for notions of linear approximations, of objects on manifolds, and of maps between manifolds. Recall that we already have notions of derivatives of maps \( \Phi: M \to N \), but only with respect to coordinate charts \((u, U)\) and \((v, V)\), in terms of the maps \((v \circ \Phi \circ u^{-1}): \mathbb{R}^m \to \mathbb{R}^n\). Clearly a coordinate-free notion is desirable.
- Vector fields are intimately connected to differential equations / dynamical systems. Thus tangent vectors should provide a means to describe dynamical systems on manifolds.
- Some vector fields are gradient vector fields, i.e. are the derivatives of some potential function. Said differently, vector fields should generalize partial differential equations for unknown functions. \([\text{Aside: This will lead to a notion of co-tangent vector fields.}]\)
- We defined arc-length as an integral of the speed. In general tangent vectors may provide a means on which to base a generalized notion of distance. \([\text{This will lead to Riemannian metrics in the second half of this course.}]\)
- We defined curvature for curves in terms of the rate of change of the tangent vectors. Thus we expect that a general notion of (comparing) tangent spaces (at different points) should provide for a notion of curvature. \([\text{This raises the question of which of the two comes first, the notion of curvature, or the means to compare tangent spaces at different points.}]\)

Should our definition allow tangent spaces at different points of a manifold to have nonempty intersection? E.g. consider the unit-circle \( S^1 \) imbedded in the plane \( \mathbb{R}^2 \). The tangent lines to \( S^1 \) at \( p = (1, 0) \) and at \( q = (0, 1) \) intersect nontrivially at \((1,1)\). On the other hand, if we think of the tangent vectors \( v_p = (0, 1) \) and \( v_q = (1,0) \) as arrows based at \( p \) and \( q \) respectively, then we certainly think of them as different.

This brings up a larger issue of distinguishing vectors (arrows) that may be moved around and vectors that are rooted at a fixed point. There are many applications where it is advantageous to consider equivalence classes of directed line-segments, equivalence meaning that they may be transformed into each other by parallel translation. On the other hand, there are many places where it is appropriate to consider vectors that are rooted, or fixed at their base points (e.g. velocity vectors to a curve).

We shall use the next section as an opportunity to bring clarity to these issues and make very precise definitions (which may always be relaxed where this causes no trouble).
3.2 Tangent spaces

There are many different ways in which one may motivate an eventual construction of tangent spaces to a general manifold. One typically starts from surfaces in Euclidean spaces, then considers more abstractly immersed manifolds in higher dimensional Euclidean spaces, and eventually tries to develop a notion that works in abstract settings, yet reduces to the familiar ones in Euclidean settings. For a lengthy such discussion see Spivak Vol. I ch. 3.

An intuitive (and very useful) way to define tangent vectors to a manifold $M$ at a point $p$ is as equivalence classes of curves. Roughly, two curves are equivalent if they have the same velocity vector at $p$—but this would be circular as we don’t have notions of velocity vectors for general curves on manifolds. So the next best thing is to declare any two smooth curves $\sigma, \gamma: (-\varepsilon, \varepsilon) \to M$ (with $\sigma(0) = \gamma(0) = p$) equivalent if $\frac{d}{dt}|_{t=0}(f \circ \sigma) = \frac{d}{dt}|_{t=0}(f \circ \gamma)$ for all smooth functions $f \in C^\infty(p)$ (that is, smooth functions defined on some neighborhood of $p$).

**Exercise 3.1** Further explore how this leads to a notion of tangent spaces that is basically the same as the one we define below. In particular, equip the collection of equivalence classes with an addition and scalar multiplication (make sure that these are well-defined). Check that the space of tangent vectors at a point is indeed an $m$-dimensional vector space. In a coordinate chart find a basis for the tangent space (e.g. provide representatives (curves) for $m$ equivalence classes that form a basis). Show how to write any tangent vector as a linear combination of this basis. Analyze how the coordinates of a tangent vector transform under local coordinate changes on the manifold. Match this notion of tangent spaces to the one provided below.

The exercise already hinted at a useful connection between tangent vectors and derivatives. Indeed, going back to Euclidean spaces, say e.g. $M = \mathbb{R}^2$ consider the relation between a vector $\tilde{u} = (u^1, u^2) = u^1 \partial_1 + u^2 \partial_2$ and the directional derivative (operator) $D_{\tilde{u}}|_p$, defined by $D_{\tilde{u}}|_p f = u^1 \cdot (D_1 f)(p) + u^2 \cdot (D_2 f)(p)$, commonly also written as $D_{\tilde{u}} f(p) = \langle u, \nabla f \rangle (p)$.

Clearly any vector $\tilde{u}$ uniquely determines a (directional) derivative operator $D_{\tilde{u}}|_p$. Conversely, one can easily recover the vector $\tilde{u}$ from the directional derivative $D_{\tilde{u}}$ by simply evaluating the latter on suitable functions: For example, evaluating $D_{\tilde{u}}|_p$ on the coordinate functions $f_1(x^1, x^2) = x^1$ and $f_2(x^1, x^2) = x^2$ immediately recovers the coordinates of $\tilde{u}$ as $u^1 = (D_{\tilde{u}} x^1)(p)$ and $u^2 = (D_{\tilde{u}} x^2)(p)$.

Thus there appears to be no harm in identifying the vector $\tilde{u}$ with the directional derivative operator $D_{\tilde{u}}(\cdot)(p)$. As we shall soon see, this idea will prove most beneficial since operators on spaces of functions are automatically endowed with a rich algebraic structure which is ready for us to use! Following Chevalley, we define

**Definition 3.1** A tangent vector to a manifold $M$ at a point $p \in M$ is a function $X_p: C^\infty(p) \to \mathbb{R}$ which is linear over $\mathbb{R}$ and is a derivation \footnote{A derivation is a map $D: \mathbb{R}[x] \to \mathbb{R}$ that is linear, $D(1) = 0$, and satisfies the Leibniz rule: $D(fg) = fD(g) + gD(f)$ for all $f, g \in \mathbb{R}[x]$.} on $C^\infty(p)$, i.e. which satisfies

- $X_p(\lambda f + g) = \lambda X_p(f) + X_p(g)$ for $\lambda \in \mathbb{R}$ and $f, g \in C^\infty(p)$ on their common domain.
- $X_p(f \cdot g) = X_p(f) \cdot g(p) + f(p) \cdot X_p(g)$ for all $f, g \in C^\infty(p)$ on their common domain.

The set of all tangent vectors to $M$ at $p$ is called the tangent space to $M$ at $p$, denoted $T_p M$.

Here $f \in C^\infty(p)$ means that there exists an open neighborhood $U$ of $p$ (depending on $f$) such that $f \in C^\infty(U, \mathbb{R})$. (This is a special case of the earlier definition of $C^\infty(A)$ for subsets $A \subseteq \mathbb{R}^n$ that are not necessarily open.)
Observations:

- $T_pM$ is a vector space: In particular, if $X_p, Y_p \in T_pM$ and $\lambda \in \mathbb{R}$ then $(X_p + \lambda Y_p) \in T_pM$. The addition and scalar multiplication are inherited from pointwise evaluations, i.e. $(X_p + \lambda Y_p)f = (X_pf) + \lambda(Y_pf)$.

- Tangent vectors are local operators: If $f, g \in C^\infty(M)$ agree on some neighborhood $U$ of $p$, i.e. $f\big|_U \equiv g\big|_U$, then $X_pf = X_pg$ for all $X_p \in T_pM$.

  Technically, the “natural domain” of tangent vectors $X_p \in T_pM$ are germs of functions at $p$: The germ of a function $f \in C^\infty(p)$ is defined as the set of all $g \in C^\infty(p)$ for which there exists an open neighborhood $U$ of $p$ so that $f\big|_U \equiv g\big|_U$.

- If $M = \mathbb{R}^m$ then the tangent vectors to $M$ at a point $p$ are precisely the directional derivatives evaluated at $p$. (One direction is obvious. For the other see the calculation below for a general manifold.)

- If $(u, U)$ is a chart about $p \in M$, then $\frac{\partial}{\partial u^i}\big|_p \in T_pM$ for $i = 1, \ldots, m$. Thus also $\sum_{j=1}^m a_j \frac{\partial}{\partial u^j}\big|_p \in T_pM$ for all $a^j \in \mathbb{R}$. To verify this assertion, recall the definition

  $$\frac{\partial f}{\partial u^i}\big|_p = D_i(f \circ u^{-1})\big|_{u(p)}$$

  and use the familiar properties of partial derivatives in $\mathbb{R}^n$ to manipulate e.g.

  $$\frac{\partial(f \cdot g)}{\partial u^i}\big|_p = D_i((f \cdot g) \circ u^{-1})\big|_{u(p)}$$

  $$= D_i((f \circ u^{-1}) \cdot (g \circ u^{-1}))\big|_{u(p)}$$

  $$= D_i(f \circ u^{-1})\big|_{u(p)} \cdot (g \circ u^{-1})(u(p)) + (f \circ u^{-1})(u(p)) \cdot D_i(g \circ u^{-1})\big|_{u(p)}$$

  $$= \frac{\partial f}{\partial u^i}\big|_p \cdot g(p) + f(p) \cdot \frac{\partial g}{\partial u^i}\big|_p$$

  More interesting is the converse, i.e. that in any chart $(u, U)$ about $p \in M$ every tangent vector $X_p \in T_pM$ may be expressed as a linear combination of the partial derivatives $\frac{\partial}{\partial u^i}$, $i = 1, \ldots, m$:

  **Theorem 3.1** If $(u, U)$ is a chart about $p \in M^m$ then $\left\{ \frac{\partial}{\partial u^1}\big|_p, \ldots, \frac{\partial}{\partial u^m}\big|_p \right\}$ is a basis for $T_pM$.

  **Corollary 3.2** If $(u, U)$ is a chart about $p \in M^m$ and $X_p \in T_pM$ then $X_p = \sum_{i=1}^m (X_pu^i) \cdot \frac{\partial}{\partial u^i}\big|_p$.

  **Corollary 3.3** If $(u, U)$ and $(v, V)$ are charts about $p \in M^m$ then $\frac{\partial}{\partial v^i}\big|_p = \sum_{j=1}^m \frac{\partial u^j}{\partial v^i}\big|_p \cdot \frac{\partial}{\partial u^j}\big|_p$.

Consider this last corollary as a statement about (linear) bases changes in the tangent space associated to (nonlinear) changes of local coordinates on the manifold.

Before proving this theorem we establish two lemmas, one rather obvious, and the other using a construction that is quite useful in a variety of places.

**Lemma 3.4** If $c : M \to \mathbb{R}$ is a constant function and $X_p \in T_pM$ then $(X_pc) = 0$

**Proof.** Use the linearity and Leibniz rule for differentiating products to establish

$$c \cdot X_p(1) = X_p(c \cdot 1) = (X_pc) \cdot 1 + c \cdot (X_p1)$$

and hence $(X_pc) = 0$. □
Lemma 3.5 Let \((u, U)\) be a chart about \(p \in M^m\) with \(u(p) = x_0\) and \(f \in C^\infty(p)\). Then there exist \(f_i \in C^\infty(p)\) such that

- \(f_i(p) = \frac{\partial f}{\partial x^i} \bigg|_p\) for \(i = 1, \ldots, m\), and

- \(f(q) = f(p) + \sum_{i=1}^m (u^i(q) - u^i(p)) f_i(q)\)

Compare this assertion to a first-order Taylor approximation. Here we have equality (as opposed to an approximation) – but the functions \(f_i\) are evaluated at the variable point \(q\) as opposed to fixed derivatives evaluated at the fixed point \(p\) in the Taylor approximation.

Proof (of the lemma). Using the local coordinates we reduce the proof to Euclidean spaces: Write \(x = u(q)\) and \(x_0 = u(p)\), rewrite the second statement of the lemma as

\[
f(u^{-1}(x)) = f(u^{-1}(x_0)) + \sum_{i=1}^m \left( u^i(u^{-1}(x)) - u^i(u^{-1}(x_0)) \right) f_i(u^{-1}(x))
\]

Write \(g\) for \(f \circ u^{-1} : u(U) \subseteq \mathbb{R}^m \to \mathbb{R}\). The desired functions \(g_j : u(U) \subseteq \mathbb{R}^m \to \mathbb{R}\) are such that

\[
g(x) = g(x_0) + \sum_{i=1}^m (x^i - x^i_0) g_i(x).
\]

After shrinking the neighborhood \(U\), if necessary, we may assume that \(u(U) \subseteq \mathbb{R}^m\) is star-shaped with respect to \(x_0\), i.e. for every \(x \in u(U)\) the line segment \(\{x_0 + t \cdot (x - x_0) : t \in [0, 1]\}\) is contained in \(u(U)\). For any fixed \(x \in u(U)\) consider the curve \(\sigma_x : [0, 1] \to u(U)\) defined by \(\sigma(t) = x_0 + t \cdot (x - x_0)\). Via the fundamental theorem of calculus and the chain rule

\[
g(x) = g(x_0) + \sum_{j=1}^m (x^j - x^j_0) \left[ \int_0^1 (D_j g)(\sigma_x(t)) \, dt \right]
\]

Note the constant derivative \(\sigma'_x = (x - x_0)\) for the curve in \(\mathbb{R}^m\) – this makes no sense on a general manifold, but is coordinate dependent. One immediately verifies that with this definition \(g_j(x_0) = (D_j g)(x_0)\) (in this case \(\sigma_x\) is a constant curve). Consequently \(f_j(p) = g_j(u(p)) = (D_j g)(u(p)) = (D_j (f \circ u^{-1}))|_{u(p)} = \frac{\partial f}{\partial x^j} \bigg|_p\). Since \(g \in C^\infty\), also \(g_j \in C^\infty\) and \(f_j \in C^\infty\).

Proof (of the theorem). Suppose \((u, U)\) is a chart about \(p \in M^m\), \(X_p \in T_p M\) and \(f \in C^\infty(M)\).

Using lemma 3.5 there exist (on some open neighborhood of \(p\)) suitable functions \(f_i\) such that we may rewrite \(X_p f\) as \(X_p f = X_p \left( f(p) + \sum_{j=1}^m (u^j(p) - u^j(0)) f_j \right)\). Using the linearity of \(X_p\), the Leibniz rule, and that \(X_p c = 0\) for any constant function this yields:

\[
X_p f = f(p) + \sum_{j=1}^m \left( X_p (u^j - 0) f_j(p) + (u^j(p) - u^j(0)) \cdot (X_p f) \right),
\]

i.e. \((X_p f) = \sum_{j=1}^m (X_p u^j) \cdot f_j(p)\). By lemma 3.5 this is equal to \((X_p f) = \sum_{j=1}^m (X_p u^j) \frac{\partial f}{\partial u^j} \bigg|_p\).

Since this holds for all \(f \in C^\infty(p)\), we conclude \(X_p = \sum_{j=1}^m (X_p u^j) \frac{\partial}{\partial u^j} \bigg|_p\).
3.3 Tangent maps, part I

For every smooth map $F: \mathbb{R}^n \to \mathbb{R}^m$ between Euclidean spaces and any point $x \in \mathbb{R}^n$ the derivative of $F$ at $x$ is a linear map $(DF)(x)(\cdot): \mathbb{R}^n \to \mathbb{R}^m$. Now that we have tangent spaces to manifolds, we are ready to associate analogous (linear) tangent maps (between tangent spaces) to smooth maps (between manifolds).

**Definition 3.2** Suppose $\Phi: M \to N$ is a smooth map between manifolds and $p \in M$. The tangent map $\Phi_p: T_pM \to T_{\Phi(p)}N$ (of $\Phi$ at $p$) is defined for $X_p \in T_pM$ and $f \in C^\infty(\Phi(p))$ by

$$(\Phi_p X_p) f = X_p(f \circ \Phi)$$

(38)

How else could $\Phi_p$ be defined? It is immediate that if $\Phi = \text{id}$ then $\Phi_p = \text{id}_{T_pM}$. Also, it follows immediately from the definition that

**Proposition 3.6** Suppose $\Phi: M^m \to N^n$ and $\Psi: N^m \to P$ are smooth maps between manifolds, and $p \in M$. Then (note the preservation of the order of $\Phi$ and $\Psi$)

$$(\Psi \circ \Phi)_p = \Psi_{\Phi(p)} \circ \Phi_p$$

(39)

**Exercise 3.2** Prove the proposition 3.6.

In local coordinates the tangent map is given by matrix multiplication. More specifically, suppose $(u,U)$ and $(v,V)$ are charts about $p \in M^m$ and $\Phi(p) \in N^n$, respectively, and $f \in C^\infty(\Phi(p))$. From the definitions calculate (using the chain-rule)

$$\left(\Phi_p \frac{\partial}{\partial u^i}\right)_p f = \frac{\partial}{\partial v^j}\big|_p (f \circ \Phi)$$

$$= D_j(f \circ \Phi \circ u^{-1})|_{u(p)}$$

$$= D_j((f \circ v^{-1}) \circ (v \circ \Phi \circ u^{-1}))|_{u(p)}$$

$$= \sum_{i=1}^n D_i((f \circ v^{-1})|_{v(p)} \cdot D_j(v^i \circ \Phi \circ u^{-1})|_{u(p)}$$

$$= \left(\sum_{i=1}^n \frac{\partial (v^i \circ \Phi)}{\partial u^i}\big|_{u(p)} \cdot \frac{\partial}{\partial v^j}\big|_{\Phi(p)} f\right)$$

In concrete examples, using local coordinates, it is convenient to express tangent vectors as column vectors. E.g. suppose, as before, $\Phi \in C^\infty(M^m,N^n)$, and $(u,U)$ and $(v,V)$ are charts about $p \in M^m$ and $\Phi(p) \in N^n$. If $X_p \in T_pM$ is a tangent vector at $p$, let $a = (a^1,\ldots,a^m)^T$ be the column vector with components $a^i = X_p u^i$, representing $X_p = \sum_{i=1}^m a^i \frac{\partial}{\partial u^i}\big|_p$. Similarly, let $b = (b^1,\ldots,b^n)^T$ be the column vector with components $b^j = (\Phi_p X_p) v^j$, representing the image $\Phi_p X_p \in T_{\Phi(p)}N$. These column vectors $a$ and $b$ are related by matrix multiplication $b = Ca$ where $C$ is the $(n \times m)$ matrix with components $C_{ij} = \frac{\partial (v^i \circ \Phi)}{\partial u^j}\big|_{\Phi(p)}$.

Formally, one may go further, and write $\alpha$ for the row-vector with components $\alpha_i = \Phi_p \frac{\partial}{\partial u^i}\big|_{p}$ and $\beta$ for the row-vector with components $\alpha_i = \frac{\partial}{\partial v^i}\big|_{\Phi(p)}$. Then formally, the images of the basis vectors $\frac{\partial}{\partial u^i}\big|_p$ are obtained by right matrix multiplication, i.e. $\alpha = \beta \cdot C$. This matches with the observation that formally $\Phi_p(X_p) = \alpha a = (\beta C)a = \beta(Ca) = \beta b$. While this is all simple (formal) matrix algebra, it is worthwhile to remember that when transforming formal vectors of basis elements these are multiplied by the transformation matrix in a way opposite to the multiplication familiar for transforming specific vectors.
**Exercise 3.3** Suppose $\Phi \in C^\infty(M^m, N^n)$ and $\Psi \in C^\infty(N^n, P^r)$ are smooth maps between manifolds, $p \in M$ and $f \in C^\infty(P)$. Furthermore, suppose $(u,U)$, $(v,V)$ and $(w,W)$ are local coordinate charts about $p \in M$, $\Phi(p) \in N$ and $(\Psi \circ \Phi)(p) \in P$, respectively. Verify that the matrix representing $(\Psi \circ \Phi)_*p$ with respect to $(u,U)$ and $(w,W)$ is the (matrix-)product of the matrices representing $\Phi_*p$ (with respect to $(u,U)$ and $(v,V)$) and $\Psi_*\Phi(p)$ (with respect to $(v,V)$ and $(w,W)$).

We digress a little to consider tangent spaces of immersed manifolds in $\mathbb{R}^n$ which justly familiar pictures of tangent planes. Suppose that $\Phi \in C^\infty(M^m, \mathbb{R}^n)$ is an immersion at $p \in M$, i.e. $\text{rank}_p \Phi = m$. Using local coordinates $(u,U)$ about $p \in M^m$ and the standard coordinates $(x, \mathbb{R}^n)$ in the range, the rank condition says that the $(n \times m)$-matrix with components $\frac{\partial \Phi}{\partial u_i}\bigg|_p$ has rank $m$, and $\Phi_*p$ is a monomorphism (a linear one-to-one map) from $T_pM$ to $T_{\Phi(p)}\mathbb{R}^n$.

The tangent vectors $\Phi_*p \left( \frac{\partial}{\partial u_i}\right) \in \Phi_*p(T_pM) \subseteq T_{\Phi(p)}\mathbb{R}^n$ are linearly independent, and span an $m$-dimensional subspace of $T_{\Phi(p)}\mathbb{R}^n$ which is usually pictured as a tangent line/plane, . . . . The image of any tangent vector $X_p \in T_pM$ in the standard coordinates may be written as $\Phi_*pX_p = \sum_{i=1}^n b_i D_i|_{\Phi(p)}$. Now, if $f \in C^\infty(\mathbb{R}^n)$ is a function such that $f \circ \Phi \equiv 0$ in a neighborhood of $p \in M$, then

$$0 \equiv X_p(f \circ \Phi) = (\Phi_*pX_p)f = \sum_{i=1}^n b_i (D_i f)|_{\Phi(p)}$$

which in calculus notation looks like $\Phi_*pX_p \perp (\text{grad } f)(\Phi(p))$, or $0 = \langle b, (\nabla f)(\Phi(p)) \rangle$. Recall if $(\text{grad } f)(\Phi(p)) \neq 0$, then $f^{-1}(0)$ is (locally) a smooth hypersurface in $\mathbb{R}^n$ near $\Phi(p)$ and thus $(\Phi_*pX_p)$ may be pictured as lying in the tangent hyperplane to the hypersurface $f^{-1}(0)$ at $\Phi(p)$.

**Exercise 3.4** Generalize this discussion to the case when there are functions $f_1, \ldots, f_n \in C^\infty(\Phi(p))$ with $f^i \cdot \Phi \equiv 0$ and linearly independent gradients $(\text{grad } f^i)(\Phi(p))$.

**Example 3.1** As a hands-on example consider an immersion of the Moebius-strip into $\mathbb{R}^3$. One way to represent the Moebius-strip is as the quotient $M = \mathbb{R}/\sim$ of the rectangle $R = [0, 2\pi] \times (-1,1)$ two of whose edges have been identified by $(0,t) \sim (2\pi,-t)$. Define a map $\Phi: M = \mathbb{R} \rightarrow \mathbb{R}^3$ by $\Phi(\Theta,t) = \left( (2 + t \cos(\frac{\Theta}{2})) \cos \Theta, \ (2 + t \cos(\frac{\Theta}{2})) \sin \Theta, \ t \sin(\frac{\Theta}{2}) \right)$.

**Exercise 3.5** Explicitly calculate the images $\Phi_*p \left( \frac{\partial}{\partial \Theta}\right) \bigg|_p$ and $\Phi_*p \left( \frac{\partial}{\partial t}\right) \bigg|_p$ at any point $p = (\Theta,t) \in U = (0,2\pi) \times (-1,1) \subseteq M$. Verify that $\Phi$ is indeed an immersion.

The image $\Phi(M) \subseteq \mathbb{R}^3$ may be thought of as a surface in the usual way. The images of the tangent vectors to $M$ calculated in the exercise may be thought of as tangent vectors in the usual sense. It is easily seen that the map $\Phi$ is indeed well-defined on $M$ (as opposed to only on the rectangle $R$) because $\Phi(0,t) = \Phi(2\pi,-t)$. For $p \in U$ the map $\Phi_*p$ is well-defined, but problems arise when trying to extend $\Phi$ and $\Phi_*q$ continuously to all $q \in M$. Indeed, we should consider $((\Theta,t),U)$ as a local coordinate chart of $M$. In the language of the next sections $\Phi_*$ maps the coordinate vector fields $\frac{\partial}{\partial \Theta} = \frac{\partial}{\partial \Theta}$ and $\frac{\partial}{\partial t} = \frac{\partial}{\partial t}$ (which are only defined on $U$) to vector fields on $\Phi(U)$, but the vector field $\frac{\partial}{\partial \Theta}$ cannot be continuously extended to a vector field on all of $M$. 

Classnotes for Introduction to Differential Geometry. Matthias Kawski. February 28, 2003 38
We conclude this first section on tangent maps with a generalization of our earlier local submersion theorem.

**Theorem 3.7** Suppose that \( \Phi \in C^\infty(M^m, N^n) \) is a smooth map between manifolds and \( P^r \subseteq N \) is a smooth submanifold. If

\[
\Phi_p(T_pM) + T_{\Phi(p)}P = T_{\Phi(p)}N \quad \text{for every} \quad p \in \Phi^{-1}(P)
\]

then \( \Phi^{-1}(P) \subseteq M \) is a submanifold of \( M \) of dimension \((m - (n - r))\).

In general one calls a smooth map \( \Phi \in C^\infty(M, N) \) between manifolds *transverse to* (a submanifold) \( P \subseteq N \) along (a submanifold) \( L \subseteq M \) if \( \Phi_*(T_pM) + T_{\Phi(p)}P = T_{\Phi(p)}N \) for all \( p \in L \cap \Phi^{-1}(P) \).

The theorem motivates the notion of codimensions as opposed to dimensions of submanifolds. More specifically, for an \( r \)-dimensional submanifold \( P^r \subseteq N^n \) of an \( n \)-dimensional manifold \( N \) the codimension of \( P \) (in \( N \)) is defined as \( (n - r) \). The theorem then simply states that if \( P \subseteq N \) is a submanifold of codimension \( k \) and \( \Phi \) is transversal to \( P \) (along \( M \)) then \( \Phi^{-1}(P) \subseteq M \) is a submanifold of the same codimension \( k \).

**Proof.** Suppose that \( \Phi \in C^\infty(M^m, N^n) \) and \( p \in \Phi^{-1}(P) \subseteq M \) is in the preimage of a smooth submanifold \( P^r \subseteq N \). Using theorem 2.13 choose an adapted chart \((v, V)\) about \( \Phi(p) \in N \) such that \( v(\Phi(p)) = 0 \) and such that the restriction of \( v = (v^1, \ldots, v^r) \) to the set \( W = \{ q \in N : v^{r+1}(q) = \ldots = v^n(q) = 0 \} \) is a chart of \( P \) about \( \Phi(p) \).

Define \( \Psi : V \mapsto \mathbb{R}^{n - r} \) by \( \Psi = (v^{r+1}, \ldots, v^n) \). Let \( U = \Phi^{-1}(V) \subseteq M \).

Then \( p \in \Phi^{-1} \circ \Psi^{-1}(0) = (\Psi \circ \Phi)^{-1}(0) \) and \( (\Psi \circ \Phi)_p = \Psi \circ \Phi_\ast : T_pM \mapsto T_{\Phi(p)}N \).

Use that \( D(\Psi \circ v^{-1}) = (0, I_{n-r}) \) and that the kernel of \( \Psi \circ \Phi(p) \) is precisely \( T_{\Phi(p)}N \). Together with \( \Phi_\ast(T_pM) + T_{\Phi(p)}P = T_{\Phi(p)}N \) this establishes that the restriction of \( \Psi \circ \Phi(p) \) to the image of \( \Phi_\ast \) (i.e. to \( \Phi_\ast(T_pM) \)) has full rank, and hence \( \text{rank}(\psi \circ \Phi)_p = n - p \) (using corollary 2.14). \( \blacksquare \)

**Exercise 3.6** Revisit the Hopf map \( \Phi : S^3 \mapsto S^2 \) (of exercise 2.30), i.e., the restriction (to \( S^3 \)) of \( \tilde{\Phi} : \mathbb{R}^4 \mapsto \mathbb{R}^3 \) defined by \( \tilde{\Phi}(a) = (2a_1 a_3 + 2a_2 a_4, 2a_2 a_3 - 2a_1 a_4, a_1^2 + a_2^2 - a_3^2 - a_4^2) \).

Considering the usual imbeddings of the spheres into \( \mathbb{R}^4 \) and \( \mathbb{R}^3 \), respectively, describe the preimages \( \Phi^{-1}(P_c) \) of the meridians \( P_c = S^2 \cap \{ x \in \mathbb{R}^3 : x^3 = c \} \) for \(-1 \leq c \leq 1\).
3.4 The tangent bundle

It is natural to assemble all tangent spaces of a manifold together into a new structure – and conceivably, this set should again have a natural manifold structure. We will omit some of the more technical details of this construction and refer to e.g. Spivak vol.I ch.3, especially exercise 1. As discussed earlier, it is desirable to distinguish between tangent vectors at different points, and we define:

Definition 3.3 The tangent bundle (as a set) $TM$ to a manifold $M$ is the (disjoint) union of all tangent spaces to $M$ at all points $p \in M$.

$$TM = \left\{ (p, X_p) \in M \times \bigcup_{p \in M} T_pM : X_p \in T_pM \right\}$$ (42)

The bundle projection $\pi: TM \hookrightarrow M$ is defined by $\pi(p, X_p) = p$. The fiber over $p \in M$ is the preimage $\pi^{-1}(p) = \{p\} \times T_pM$. A section, or tangent vector field is a map $X: M \hookrightarrow TM$ that satisfies $\pi \circ X = id_M$.

Vector fields are basically functions that assign to each $p \in M$ a tangent vector in $T_pM$. But occasionally it is convenient to identify the fiber $\pi^{-1}(p)$ with $T_pM$ – technically this includes a tacit projection onto the second factor. Practically, this means writing $X_p$ where more precise notation is $X(p) = (p, X_p)$ for the pair that consists of the point $p \in M$ and the tangent vector $X_p \in T_pM$.

To illustrate the natural manifold character of tangent bundles, consider the example of $M = S^1$. The naive collection of all tangent lines to the imbedded circle $S^1 \subseteq \mathbb{R}^2$ is full of intersections. More suitable for our purposes is to imbed the circle in $\mathbb{R}^3$ as $\tilde{S}^1 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1, \ x_3 = 0\}$ and attach at every point $p \in \tilde{S}^1$ a vertical (!) line, yielding a cylinder. As a set, this cylinder is in bijection with the (disjoint) collection of all tangent lines to the circle imbedded in the plane. It is clear that one can consistently choose an orientation of the lines, and even more a consistent scaling. Intuitively identify the naive tangent vector $\left( (\cos \Theta, \sin \Theta), (-L \sin \Theta, L \cos \Theta) \right)$ with the point $(\cos \Theta, \sin \Theta, L) \in \mathbb{R}^3$.

Within this picture, a vector field on the circle may be visualized as the graph of a function $\Theta \longmapsto (\cos \Theta, \sin \Theta) \longmapsto L(\theta)$. If the vector field is continuous and nonvanishing, then the graph lies entirely above, or entirely below the plane $x_3 = 0$.

In complete analogy, we may intuitively think of the tangent bundle $T\mathbb{R}$ of the real line $\mathbb{R}$ as the plane $\mathbb{R}^2$. However, for dimensional reasons it is clear that these two examples are the only tangent bundles amenable to such immediate visualization. How quickly things get complicated becomes clear if one tries to think of $TS^2$ as a sphere with a (different) plane attached to each of its points. A vector field on the sphere simply selects one point on each plane. However, from algebraic topology it is known that there does not exist any continuous vector field on the sphere that vanishes nowhere. In our picture this means that it is impossible to continuously select one point on each tangent plane avoiding the origin (zero-vector) in each $T_pS^2$. Intuitively $TS^2$ must be a nontrivially twisted, (when compared to e.g. $TS^1$ which is the very tame cylinder), i.e. it must be very different from the trivial Cartesian product $S^2 \times \mathbb{R}^2$.

We proceed more abstractly to endow $TM$ for a general smooth manifold $M$ with a manifold structure. The key idea is that locally, above a chart $(u, U)$ (which itself is homeomorphic to
exists functions $w$.

Suppose $(u,U)$ is a chart of $M$ about $p$. Consider the subset $\bar{U} = \pi^{-1}(U) \subseteq TM$. Every point $Q \in \bar{U}$ is a pair $Q = (q,X_q)$ with $X_q \in T_qM$. Since $\{\frac{\partial}{\partial \omega^i}|_q : 1 \leq j \leq m\}$ is a basis for $T_qM$ there exists functions $w^j: U \mapsto \mathbb{R}$ (indeed, $w^j(Q) = X_q w^j$) such that

$$Q = \left( \pi(Q), \sum_{j=1}^{m} w^j(Q) \frac{\partial}{\partial \omega^j}|_{\pi(Q)} \right).$$

(43)

It is natural to define a map $\bar{u}: \bar{U} \mapsto \mathbb{R}^m$ by

$$\bar{u} = \left( u^1 \circ \pi, \ldots, u^m \circ \pi, w^1, \ldots, w^m \right).$$

(44)

It is clear that $\bar{u}$ is a bijection from $\bar{U}$ to $\mathbb{R}^m$ (This assumes that $u$ is a bijection from $U$ to $\mathbb{R}^m$, i.e. $(u,U)$ a chart in the original sense. Alternatively, $\bar{u}$ is a bijection from $\bar{U}$ onto $u(U) \times \mathbb{R}^m \subseteq \mathbb{R}^{2m}$.)

We want for each chart $(u,U)$ in the $C^\infty$-differentiable structure of $M$ the associated map $\bar{u}$ to be a homeomorphism. One may simply endow $TM$ with the weakest (i.e. coarsest) topology in which all maps $\bar{u}$ are continuous. More constructively, we consider the collection $\mathcal{T}$ of all subsets $O \subseteq TM$ which are such that for every point $Q \in TM$ and every chart $(u,U)$ of $M$ for which $\pi(Q) \in U$, there exists an open set $W \in \mathbb{R}^{2m}$ containing $\bar{u}(Q)$ such that $\bar{u}^{-1}(W) \subseteq O$.

**Exercise 3.7** Show that $\mathcal{T}$ is a topology on $TM$, i.e. $\emptyset, TM \in \mathcal{T}$, and that $\mathcal{T}$ is closed under finite intersections and arbitrary unions.

**Exercise 3.8** Check that when $TM$ is endowed with this topology $\mathcal{T}$ then for every chart $(u,U)$ the associated map $\bar{u}$ is a homeomorphism, i.e. is continuous with continuous inverse.

We next calculate the transition maps $(\bar{v} \circ \bar{u}^{-1}): \bar{u}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^{2m} \mapsto \bar{v}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^{2m}$ to verify that the charts are indeed $C^\infty$-related.

Thus assume that $Q \in \bar{U} \cap \bar{V} \subseteq TM$ and that $\bar{u}(Q) = (x,y) \in \mathbb{R}^m \times \mathbb{R}^m$. Calculate

$$\bar{v}(Q) = (\bar{v} \circ \bar{u}^{-1})(x,y) = \bar{v}(u^{-1}(x), \sum_{j=1}^{m} y^j \frac{\partial}{\partial \omega^j}|_{u^{-1}(x)})$$

(45)

In order to evaluate the second $m$-components of $\bar{v}(Q)$ change the basis in $T_{\pi(Q)}M$ from the $\frac{\partial}{\partial \omega^i}|_{\pi(Q)}$ to the $\frac{\partial}{\partial v^j}|_{\pi(Q)}$, yielding

$$(\bar{v} \circ \bar{u}^{-1})(x,y) = \bar{v} \left( u^{-1}(x), \sum_{j=1}^{m} y^j \sum_{i=1}^{m} \frac{\partial v^j}{\partial \omega^i} \bigg|_{u^{-1}(x)} \frac{\partial}{\partial v^i} \bigg|_{u^{-1}(x)} \right)$$

(46)
After interchanging the order of summation and regrouping one reads off the components of \( \tilde{v}(Q) \)

\[
(\tilde{v} \circ \tilde{u}^{-1})(x, y) = ((v \circ u^{-1})(x), \sum_{j=1}^{m} y_j \frac{\partial v^j}{\partial u^i}(u^{-1}(x)), \ldots, \sum_{j=1}^{m} y_j \frac{\partial v^m}{\partial u^i}(u^{-1}(x)))
\]

(47)

It is easily seen that the map \( (\tilde{v} \circ \tilde{u}^{-1}): \tilde{u}(U \cap V) \subseteq \mathbb{R}^{2m} \rightarrow \tilde{v}(U \cap V) \subseteq \mathbb{R}^{2m} \) is a smooth map: By hypothesis the first \( m \)-components are smooth maps. The second \( m \) components are linear in \( y \) and smooth functions of \( x \), and hence the combined map is smooth. \([\text{Incidentally, if working in the class of } C^r \text{-manifolds, this calculation shows that when one starts with a } C^r \text{ atlas for } M, \text{ then one obtains, as might be expected, a } C^{r-1} \text{ atlas for } TM.\]

In order for \( TM \) to qualify as a smooth manifold we still need that the topology is reasonably nice (metrizable, or equivalently that \( (TM, \mathcal{T}) \) is paracompact). For the technical details we refer to Spivak vol.I ch.3, especially exercise 1. Here we sketch only some basic ideas. Since a manifold \( M \) is locally homeomorphic to a Euclidean space and it is assumed to be metrizable (or equivalently paracompact), each connected component of \( M \) is second countable. \([\text{This means that there is a countable basis for the topology on each connected component of } M. \text{ A basis for a topology is a collection of open sets that covers the space, and such that whenever a point } x \text{ is contained in basic open sets } B_1 \text{ and } B_2, \text{ then there exists a basic open set } B_3 \text{ such that } x \in B_3 \subseteq B_1 \cap B_2.]\]

Following Spivak vol.I ch.3 ex. 1, construct a sequence of functions that \textit{separates points and closed sets}, use these to produce a sequence of bounded metrics \( d_i \) and finally piece these together e.g. via \( d = \sum_{i=1}^{\infty} 2^{-i} d_i \).

This construction of the tangent bundle shall serve as a model for similar constructions of more general \textit{vector bundles} in which the tangent spaces \( T_p M \) are replaced by other suitable vector spaces. Formally, a vector bundle is a triple \((E, B, \pi)\) (or actually, a five-tuple \((E, B, \pi, \oplus, \circ)\)) consisting of a \textit{total space} \( E \), a \textit{base space} \( B \) and a bundle projection \( \pi: E \rightarrow B \) which is a continuous surjective map. \[\text{The linear operations } \oplus \text{ and } \circ \text{ are defined on the fibres } \pi^{-1}(p) \text{ for } p \in B, \text{ making each fibre a vector space.} \text{ The most important condition is that a vector bundle must be \textit{locally trivial}, i.e. every } p \in B \text{ has an open neighborhood } U \text{ together with a homomorphism } \beta: \pi^{-1}(U) \rightarrow U \times \mathbb{R}^m \text{ such that for each } q \in U \text{ the restriction } \beta_{\pi^{-1}(q)} \text{ is a vector space isomorphism from the fibre } \pi^{-1}(q) \text{ to } \{q\} \times \mathbb{R}^m.\]

We already have seen the Moebius strip which is an example of a nontrivial line-bundle over the circle \( S^1 \). An upcoming section will introduce the \textit{cotangent bundle} in which the fibre are the spaces of linear functionals on the corresponding tangent spaces.

In some places it is convenient to work with functions that assign to each point \( p \in M \) a pair, triple, or \( m \)-tuple of \((co)-tangent\) vectors. These may be thought off as sections of bundles in which each fibre is a product of two, three, or \( n \) copies of the \((co)tangent\) space.

Beyond vector bundles are fibre bundles in which the fibres need not necessarily be vector spaces. Arguably the most important such is \( n \)\textit{the principal bundle} in which each fibre is a copy of the general linear group \( GL(m, \mathbb{R}) \) (the space of all invertible linear maps from \( \mathbb{R}^m \) to \( \mathbb{R}^m \)). Its distinguishing feature is that each section \( L: M \rightarrow P \text{ acts} \) on e.g. local coordinates via composition: If \((u, U)\) is a chart then \((L \circ u, U)\) is another chart (to be read as \((L \circ u: q \mapsto L_q \circ u(q), U)\) where \(L_q: \mathbb{R}^m \rightarrow \mathbb{R}^m\) is a linear map.)

To be added: Use tangent bundles for a geometric definition of \textit{orientability} for a manifold \( M \), or of vector bundle - as opposed to the purely algebraic condition in terms of charts, whether there exists atlas \( \mathcal{A} \) such that \( \det(D(v \circ u^{-1}) > 0 \) for all \((u, U), (v, V) \in \mathcal{A} \).
We digress with a brief discussion of the (lack of) trivality of the tangent bundles of spheres and its consequences. Consider the usual imbeddings of the spheres $S^m \hookrightarrow \mathbb{R}^{m+1}$ and use the standard coordinates in $\mathbb{R}^{m+1}$. Note that the tangent bundles $TS^m$ are diffeomorphic to the subspaces $\{(a, b) \in S^m \times \mathbb{R}^{m+1} : \langle a, b \rangle = 0\} \subseteq \mathbb{R}^{m+1} \times \mathbb{R}^{m+1}$, using the standard inner product in $\mathbb{R}^{m+1}$. [[This is completely different from asserting that $TS^m$ were trivial, or diffeomorphic to $S^m \times \mathbb{R}^m$.]]

Observe that when $m = (2k - 1)$ is odd, then $X = x_2 D_1 - x_1D_2 + x_3 D_3 - x_4 D_4 + \ldots x_{2k} D_{2k-1} - x_{2k-1} D_{2k}$ (all $D_j$ evaluated at $x$) is a global nonvanishing (tangent) vector field on $S^m$. Using tools from algebraic topology one may show that if $m = 2k$ is even, then no globally defined continuous nonvanishing vector field exists on the sphere $S^{2k}$. One of the more remarkable spheres is $S^3$ which admits three smooth vector fields that are everywhere linearly independent:

\[
\begin{align*}
X(x) &= -x_2 D_1 + x_1 D_2 - x_4 D_3 + x_3 D_4 \\
Y(x) &= -x_3 D_1 + x_4 D_2 + x_1 D_3 - x_2 D_4 \\
Z(x) &= -x_4 D_1 - x_3 D_2 + x_2 D_3 + x_1 D_4
\end{align*}
\] (48)

Mimicking (and repeating this construction, similar to the example for $S^{2k-1}$ above) one may construct from these three vector fields on $S^3$ three everywhere linearly independent vector fields on any sphere $S^{4k-1}$. However, it can be shown that on $S^{4k+1}$ any two smooth vector fields are linearly dependent at some point.

The example of the frame of the three everywhere linearly independent vector fields on $S^3$ motivates the notion of a parallelizable manifolds [[Abraham-Marsden p.218; Boothby p.219; not in Spivak]]:

**Definition 3.4** A manifold $M^m$ is called parallelizable if it admits a frame of $m$ everywhere linearly independent vector fields.

It is straightforward to see that a [[finite-dimensional, c.f. Abraham-Marsden]] manifold is parallelizable if and only if its tangent bundle is trivial. So far we have seen that Euclidean spaces $\mathbb{R}^m$, hence all coordinate charts $(u, U)$ are parallelizable. Also, every Lie group is parallelizable.

---

**Should have been done much earlier, in chapter 2:** A Lie group is a differentiable manifold $G$ with a group structure such that both the multiplication $: G \times G \to G$ defined by $(p, q) \mapsto pq$ and the inverse $G \to G$, defined by $p \mapsto p^{-1}$ are $C^\infty$ maps.

We already encountered several examples of Lie groups: the general linear groups $GL(n, \mathbb{R})$ of invertible linear maps on $\mathbb{R}^n$, the special linear groups $SL(n, \mathbb{R})$ (linear maps with determinant one), and the orthogonal groups $O(n)$ and special orthogonal groups $SO(n)$. Note that as manifolds $S^1$ is diffeomorphic to $SO(2)$. Similarly, $S^3$ is a double-cover of the projective space $P^3$ which is diffeomorphic to $SO(3)$ – thus shedding some light on this most versatile example.

**Exercise 3.9** Suppose $f \in C^\infty(\mathbb{R}^m, \mathbb{R})$. Show that the graph $\{(x, f(x)) : x \in \mathbb{R}^m\}$ is a parallelizable submanifold of $\mathbb{R}^{m+1}$. Is the same necessarily true for functions $f : \mathbb{R}^m \to \mathbb{R}^n$?

Returning to the tangent bundles of the spheres: It is known that the only parallelizable spheres are $S^1$, $S^3$, and $S^7$ [[Spivak vol.I, ch.3, ex. 19]]. It is no coincidence that these are the only dimensions in which one may endow the Euclidean spaces with some sort of multiplicative structure: In $\mathbb{R}^2$ this is the field of complex numbers, in $\mathbb{R}^4$ this yields the noncommutative, but still associative quaternions (or Hamilton numbers), and in $\mathbb{R}^8$ these are the Cayley numbers whose multiplication is even no longer associative.
3.5 Smooth vector fields and Lie products

We defined a vector field on a manifold to be a section of the tangent bundle, that is, a function \( X: M \rightarrow TM \) such that its composition \( \pi \circ X \) with the bundle projection is the identity on \( M \). Rather than considering arbitrary such functions, our interest is primarily in those that vary smoothly (in topological considerations continuity may suffice). Since a vector field is defined as a map between manifolds \( M \) and \( TM \) we already have a notion of smoothness: A vector field \( X: M \rightarrow TM \) is \( C^r \) if for every point \( p \in M \) and coordinate charts \( (u, U) \) and \( (\tilde{v}, \tilde{V}) \) about \( p \) and \( X(p) \), respectively, the map \( \tilde{v} \circ X \circ u^{-1}: \mathbb{R}^m \rightarrow \mathbb{R}^{2m} \) is a \( C^r \)-map between Euclidean spaces. For notational convenience we introduce write \( \Gamma^\infty(M) \) for the set of all smooth vector fields on \( M \).

On the other hand recall that every tangent vector \( X_p \in T_p M \) maps \( C^\infty(p) \rightarrow \mathbb{R} \). Consequently, we may view a vector field \( X \) as a mapping on the algebra \( C^\infty(M) \) of smooth functions. We expect that if \( X \) is a smooth vector field and \( f \in C^\infty(M) \) then \( (Xf) \in C^\infty(M) \). In particular, any smooth vector field is a derivation on the algebra \( C^\infty(M) \) (i.e. it satisfies \( X(fg) = (Xf)g + f(Xg) \) for all \( f, g \in C^\infty(M) \)).

**Proposition 3.8** A vector field \( X: M \rightarrow TM \) is a \( C^\infty \) vector field (i.e. \( X \in \Gamma^\infty(M) \)) if and only if for every open set \( U \subseteq M \) and every function \( f \in C^\infty(U) \) the function \( (Xf): p \mapsto X(p)f \) is again in \( C^\infty(U) \).

This proposition follows easily from the following exercise upon expanding \( (Xf)(p) \) on a chart \( (u, U) \) about \( p \) in terms of local coordinates \( (Xf)(q) = \sum_{j=1}^m (Xu^j)(q) \frac{\partial L}{\partial u^j}(q) \) and writing out the components of the map \( \tilde{u} \circ X \circ u^{-1}: \mathbb{R}^m \rightarrow \mathbb{R}^{2m} \).

**Exercise 3.10** Verify directly that a vector field \( X: M \rightarrow TM \) is \( C^\infty \) if and only if for every coordinate chart \( (u, U) \) of \( M \) the functions \( Xu^j: U \rightarrow \mathbb{R} \) are smooth.

Since every (smooth) vector field \( X \in \Gamma^\infty(M) \) maps \( C^\infty(M) \) back into itself, it is natural to consider compositions of two vector fields \( X, Y \in \Gamma^\infty(M) \). Clearly \( X \circ Y \), also written \( XY \), is again a map from \( C^\infty(M) \) into itself. However, for two functions \( f, g \in C^\infty(M) \) we calculate

\[
XY(fg) = X \left( (Yf)g + f(Yg) \right) = (XYf)g + (Yf)(Xg) + (Xf)(Yg) + f(XYg)
\]

In general there is no reason for the terms \( (Yf)(Xg) \) and \( (Xf)(Yg) \) to cancel each other, hence in general \( XY \) is not a derivation, and thus \( NOT \) a vector field! However, the commutator \( XY - YX \) clearly will be a derivation. Thus we have a product structure on \( \Gamma^\infty(M) \) — which equips this space of all smooth vector field with an important algebraic structure that invites deeper study:

**Definition 3.5** The Lie bracket or Lie product (of vector fields) is the map \([\cdot, \cdot] : \Gamma^\infty(M) \times \Gamma^\infty(M) \rightarrow \Gamma^\infty(M)\), defined by \([X, Y]f = X(Yf) - Y(Xf)\) for \( f \in C^\infty(M)\).

**Exercise 3.11** Let \( \xi = (\xi^1, \ldots, \xi^m) \) and \( \eta = (\eta^1, \ldots, \eta^m) \) be column vector fields representing two vector fields \( X, Y \in \Gamma^\infty(M) \) in a coordinate chart \( (u, U) \), i.e. \( \xi^j = (Xu^j) \) and \( \eta^j = (Yu^j) \). Verify that in these coordinates the Lie product \([X, Y]\) is represented by the column vector \((D\eta)(\frac{\partial L}{\partial u^j}) - (D\xi)\pi\) where \( D \) denotes the Jacobian matrix of partial derivatives.

**Definition 3.6** A linear vector space \( L \) equipped with a bilinear mapping \([\cdot, \cdot] : L \times L \rightarrow L\) is a Lie algebra if this map is anti-commutative and satisfies the Jacobi identity:

\[
[x, y] + [y, z] + [z, x] = 0 \quad \text{for all} \quad x, y, z \in L
\]

\[
[x, y] + [y, z] + [z, x] = 0 \quad \text{for all} \quad x, y, z \in L
\]

(50)
Exercise 3.12 Verify that \( \mathbb{R}^3 \) equipped with the standard cross-product is a Lie algebra.

Exercise 3.13 Verify that the space \( \text{so}(3) \) of skew symmetric \( 3 \times 3 \)-matrices with the product 
\[ [A, B] = AB - BA \] (matrix product) is a three dimensional Lie algebra.

Find a basis for \( \text{so}(3) \) and establish a Lie algebra isomorphism from \( \text{so}(3) \) to \( \mathbb{R}^3 \) with the cross-
product – i.e. explicitly give a bijective linear map (between vector spaces) that is also a Lie 

algebra homomorphism, meaning in this case \( \Phi([A, B]) = \Phi(A) \times \Phi(B) \) for all \( A, B \in \text{so}(3) \).

Exercise 3.14 Verify by direct calculation that the Lie product of vector fields as defined above 
equips \( \Gamma^\infty(M) \) with a Lie algebra structure. Note that this means verifying that \([ \cdot, \cdot \] \) is linear 
over \( \mathbb{R} \), i.e. \([ aX + Y, Z ] = a[X, Z] + [Y, Z] \), that it is anticommutative (obvious) and 

Exercise 3.15 Show that any associative algebra \((A, \cdot)\) is a Lie algebra under the commutator 
product defined by \([x, y] = x \cdot y - y \cdot x \).

Similarly, the set \( \mathcal{D}(A) \) of derivations on an associative algebra, that is of linear maps \( \ell : A \rightarrow A \) 
satisfying \( \ell(xy) = (\ell(x))y + x\ell(y) \) for all \( x, y \in A \) is an associative algebra 
under composition and thus a Lie algebra under the commutator as above.

Recall that if \( X, Y \in \Gamma^\infty(M) \) and \( f, g \in C^\infty(M) \) then \( fX \in \Gamma^\infty(M) \), and the usual distributive 
and mixed associative properties hold, e.g. \( (fg)X = f(gX), f(X + Y) = fX + fY, (f + g)X = 
fx + gX, 1 \cdot X = X, \ldots \) This means that \( \Gamma^\infty(M) \) is not only a vector space over \( \mathbb{R} \), but also a 
(left) \( C^\infty(M) \)-module. (It is not a vector space over \( C^\infty(M) \) since the ring of smooth functions 
is not a field.)

Given this \( C^\infty(M) \)-module structure it is natural to ask how the Lie bracket on \( \Gamma^\infty(M) \) relates 
to it. For \( X, Y \in \Gamma^\infty(M) \) and \( f, g \in C^\infty(M) \) we calculate

\[
[fX, Y]g = (fX)(Yg) - Y(fX(g)) = f(X(Yg) - Y(Xg)) - (Yf) \cdot (Xg) = \left( f[X, Y] - (Yf)X \right) g \tag{51}
\]

and conclude that the Lie bracket \([ \cdot, \cdot ] \) is not linear over \( C^\infty(M) \).

In a chart \((u, U)\) (compare also exercise 3.20) we calculate

\[
\left[ \frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^\nu} \right] f = \frac{\partial}{\partial u^\nu}(D_j(f \circ u^{-1}) \circ u) - \frac{\partial}{\partial u^\nu}(D_i(f \circ u^{-1}) \circ u) = D_j(D_i(f \circ u^{-1}) \circ u \circ u^{-1}) \circ u - \frac{\partial}{\partial u^\mu}(D_i(f \circ u^{-1}) \circ u \circ u^{-1}) \circ u = \left( D_i D_j(f \circ u^{-1}) - D_j D_i(f \circ u^{-1}) \right) \circ u \tag{52}
\]

since the mixed partial derivatives on \( C^\infty \mathbb{R}^m \) are equal. As an important corollary we obtain:

**Proposition 3.9** If \( X, Y \in \Gamma^\infty(M) \) and \( U \subseteq M \) is an open set with \( [X, Y]|_U \neq 0 \) then there does 
not exist a map \( u : U \rightarrow \mathbb{R}^m \) such that \((u, U)\) is a chart on \( M \) with \( X|_U = \frac{\partial}{\partial u^1} \) and \( Y|_U = \frac{\partial}{\partial u^2} \).

Indeed, in subsequent sections we will see that in the neighborhood of any point \( p \) at which a 
smooth vector field \( X \) does not vanish, there are always coordinates \((u, U)\) such that \( X = \frac{\partial}{\partial u^1} \).

On the other hand, generalizing the above criterion to sets of vector fields will lead to important 
Frobenius integrability theorem.
Exercise 3.16 [[This exercise is somewhat frivolous – but it is a good practice for hands-on calculations, and it hits hard at common misperceptions.]] Consider the upper half plane $M = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$ with standard rectangular coordinates $(x, y)$, with polar coordinates $(r, \theta)$ and with the mixed coordinates $(\rho, \xi)$ defined by $\rho = r$ and $\xi = x$.

- Explicitly the coordinate vector fields $\frac{\partial}{\partial r}$ and $\frac{\partial}{\partial \theta}$ as linear combinations of $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$. (In particular express the coefficients in terms of $x$ and $y$).

- Use these expressions to verify by direct calculation that $[\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}] \equiv 0$.

- Verify by direct calculation that $[\frac{\partial}{\partial r}, \frac{\partial}{\partial x}] \not\equiv 0$.

- Explain why this does not contradict that $(\xi, \rho) = (x, r)$ are admissible local coordinates.

- Calculate the $(\xi, \rho)$ coordinates of the points $(1, 0, .1)$, $(1, 1)$, and $(0, 1)$.

- Calculate $\frac{\partial}{\partial \xi}$ and $\frac{\partial}{\partial \rho}$, e.g. write these as linear combinations of $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$, and sketch these coordinate vector fields as arrows in the half-plane. Describe in words in which directions these arrows point.

- Explain the basis for possible misconceptions – e.g. in terms of partial derivatives holding fixed some variables . . .

3.6 The tangent map and vector fields

Having assembled the tangent spaces $T_pM$ at all points $p \in M$ into the tangent bundle $TM$ as a manifold, it is natural to combine the tangent maps $\Phi^*_p$ associated to a map $\Phi \in C^\infty(M, N)$ between manifolds into a map $\Phi^*: TM \mapsto TN$. This is a straightforward definition with no or few ensuing surprises. However, we shall need to look into some detail how such tangent maps interact with vector fields.

Definition 3.7 For any map $\Phi \in C^\infty(M, N)$ define an associated tangent map $\Phi_*$ by

$$\Phi_*(q, X_q) = (\Phi(q), \Phi_q(X_q)) \quad \text{for} \quad q \in M, \quad (q, X_q) \in \pi^{-1}(q)$$

(53)

Note that the tangent map $\Phi_*$ has the map $\Phi$ built in. It is straightforward to verify the following:

Proposition 3.10 If $\Phi \in C^\infty(M, N)$ and $\Psi \in C^\infty(N, P)$ then (note preservation of order)

$$(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*$$

(54)

Exercise 3.17 Show that if $\Phi \in C^\infty(M, N)$ then $\Phi_* \in C^\infty(TM, TN)$. (Use the definition of differentiability of a map between manifolds in terms of charts $((u, U))$ and $((v, V))$ for $M$ and $N$, respectively.)

It is important to note that in general there is no hope that a tangent map associated to a smooth $\Phi: M \mapsto N$ between manifold will map a vector field $X$ on $M$ to a vector field on $N$. This is immediately clear if we recall that a vector field on $N$ is a function from $N$ to $TN$. Thus if $p_1 \neq p_2 \in M$ but $\Phi(p_1) = \Phi(p_2) \in N$ then problems arise unless $\Phi_{sp_1}X_{p_1} = \Phi_{sp_2}X_{p_2}$. Similarly, if $\Phi$ is not onto, then $\Phi_*$ can at best yield a partially defined vector field on $N$. 

Classnotes for Introduction to Differential Geometry. Matthias Kawski. February 28, 2003 46
If $\Phi: M \rightarrow N$ is a diffeomorphism and $X \in \Gamma^\infty(M)$ then we define $(\Phi_\ast X): N \rightarrow TN$ by

$$(\Phi_\ast X)_q = \Phi_\ast \Phi^{-1}(q) X_{\Phi^{-1}(q)} \quad \text{for} \quad q \in N. \quad (55)$$

Sometimes this is written suggestively as $(\Phi_\ast X)_q = (\Phi_\ast X \circ \Phi^{-1})(q)$. It is clear that $\pi \circ (\Phi_\ast X) = \text{id}_N$ and that $\Phi_\ast$ is a smooth map, and hence $(\Phi_\ast X) \in \Gamma^\infty(N)$.

The most important application is when $\Phi = u: U \mapsto \mathbb{R}^m$ is a coordinate map. Indeed, we have routinely used the map $u_i$ which maps e.g. coordinate vector fields $\frac{\partial}{\partial x^i}$ to the fields $D_j$ on $\mathbb{R}^m$.

Note that it is not required for $\Phi$ to be a diffeomorphism in order for $(\Phi_\ast X)$ to make sense as a vector field on $N$ – as long as $\Phi$ is a smooth map such that $p_1 = p_2 \in M$ implies $\Phi_{x_{p_1}} X_{p_1} = \Phi_{x_{p_2}} X_{p_2}$ the definition (55) still makes sense. The following example gives a preview on how this may be used in the case that the vector field has some \textit{infinitesimal symmetries} as they will be defined in the section on Lie derivatives.

**Exercise 3.18** Consider $M = \mathbb{R}^2 \setminus \{0\}$ and $N = \mathbb{P}^1$. Let $X(x) = (ax^1 + bx^2) D_1|_x + (cx^1 + dx^2) D_2|_x$ be a linear vector field on $M$. Define a relation $\sim$ on $M$ by $x \sim y$ if there exists $\lambda \in \mathbb{R}$ such that $x = \lambda y$. Verify that $\sim$ is an equivalence relation on $M$.

Let $\Phi: M \mapsto N = \mathbb{P}^1 = \mathbb{M}/\sim$ be the canonical projection map which maps each $x \in M$ to its equivalence class $[x] = \{y \in M : y \sim x\}$. Verify that if $x \sim y$ then $\Phi_x(X_x) = \Phi_y(X_y)$ and hence we may $\Phi_\ast(X)$ does define a smooth vector field on $\mathbb{P}^1$.

Consider the local coordinate chart $(m, U)$ on $\mathbb{P}^1$ where $U = \{(x_1, x_2) : x_1 \neq 0\}$. For $m([x_1, x_2])$ is the slope of the line through the points $[(x_1, x_2)]$, i.e. $m([x_1, x_2]) = \frac{x_2}{x_1}$. Find an explicit expression for $(\Phi_\ast X)_m = f(m) \frac{\partial}{\partial m}|_m$. Interpret $(\Phi_\ast X)$ as defining a dynamical system (via $m = f(m)$) on the space of lines thru the origin. In detail discuss the special cases when $a = d$ and either $b = c = 0$ or $b = -c = -1$. In general relate the stationary points of $\Phi_\ast X$ (i.e. the zeros of $f(\theta)$) to the eigenspaces of the $2 \times 2$-matrix with entries $a, b, c$ and $d$.

**Exercise 3.19** Extend the previous exercise 3.18 to a higher dimensional case. Let $X(x) = \sum_{i,j=1}^m a_{ij} x^i D_j|_x$ be a linear vector field on $\mathbb{R}^m$. Use the coordinates $y = (y^1, \ldots, y^m)$ on the subset $U \subseteq \mathbb{P}^{m-1} = (\mathbb{R}^m \setminus \{0\})/\sim$ defined by $y^i([x]) = \frac{x^i}{x^m}$ and verify that in these coordinates $(\Phi_\ast X)$ is a quadratic vector field (representing a Riccati differential equation).

As an illustration explicitly write out the formula for $(\Phi_\ast X)$ in the case of $m = 1$. For fun explore the case where the matrix $(a_{ij})$ has a triple eigenvalue with a single Jordan block, e.g. $a_{ii} = \lambda \neq 0, a_{12} = a_{23} = 1$ and $a_{ij} = 0$ else. In particular, sketch the phase portrait for $(\Phi_\ast X)$ near $y = 0$ and relate it to the integral curves on $\mathbb{P}^2$ (or on $S^2$ which may be easier to visualize).