1. Calculus on normed vector spaces

We introduce and collect the basics of calculus on $\mathbb{R}^n$ and more generally on a normed (finite dimensional) vector space. The latter (slight) generalization will be helpful since many naturally appearing vector spaces (like the space of linear maps between $\mathbb{R}^k$ and $\mathbb{R}^l$) have no preferred way to be expressed as some $\mathbb{R}^n$. In principle, those results together with their proofs should be known, at least for $\mathbb{R}^n$.

1.1. **Normed vector spaces.**

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. A *norm* on a vector space V over the field $\mathbb{K}$ is a map $q : V \to \mathbb{R}$ with the properties

$$q(x) \geq 0 \, \text{for all} \, x \in \mathbb{K} \tag{1}$$

$$q(x) = 0 \, \text{iff} \, x = 0 \tag{2}$$

$$q(ax) = |a| q(x) \tag{3}$$

$$q(x + y) \leq q(x) + q(y) \, \text{triangle inequality} \tag{4}$$

where $|a| = \sqrt{a\bar{a}}$ for $a \in \mathbb{K}$. If a norm on a vector space is chosen we usually refer to it as a *normed vector space*. Similar to the absolute value on $\mathbb{R}$ one shows that any norm also satifies the so-called *modified triangle inequality*

$$|q(x) - q(y)| \leq q(x - y) \, .$$

Example 1.1. $V = \mathbb{R}^n$ with $q_p(x) = (\sum_{k=1}^n |x_k|^p)^{1/p}$ for $p = 1, 2, 3...$ and $q_\infty(x) = \max_{k=1,\dots,n} |x_k|$. For $p = 2$ we get the *Euclidean* norm measuring the length of x.

Example 1.2. Let V_k, $k = 1, 2$ be finite dimensional normed vector spaces over $\mathbb{K}$ with norms q_k and let $V = Hom(V_1, V_2) = \{T : V_1 \to V_2 \, ; \, T \, \mathbb{K}\text{-linear}\}$ be the vector space of linear maps between V_1 and V_2. Then

$$q(T) := \sup_{x \neq 0} \frac{q_2(Tx)}{q_1(x)} = \sup_{x, q_1(x)=1} q_2(Tx) \tag{5}$$

is a norm on V, the *operator norm.* From the definition it follows easily that one has the following *multiplicative* relation

$$q_2(Tx) \leq q(T) q_1(x) \, .$$

Example 1.3. Let V_k, $k = 1, 2$ be finite dimensional normed vector spaces over $\mathbb{K}$ with norms q_k . Then $V_1 \times V_2$ becomes a normed vector space with the norm $q(x_1, x_2) := q_1(x_1) + q_2(x_2)$

Example 1.4. There are also non-finite dimensional examples of normed vector spaces the study of which is done in *functional analysis.* The most basic example of such a vector space is the space of sequences $V = \{(x_k)_{k \in \mathbb{N}} \, ; \, x_k \in \mathbb{K}, x_k \neq 0 \text{for only finitely many k}\}$ whose elements are sequences in $\mathbb{K}$ with only finitely many sequence elements non-zero. Then the obvious analogs of the norms defined in Example 1.1 give the norms $q_p(x) = (\sum_{k=1}^\infty |x_k|^p)^{1/p}$ for $p = 0, 1, 2...$ and $q_\infty(x) = \max_{k=1,\dots,\infty} |x_k|$. Notice that there are no convergency issues since the sequences have only finitely many non-zero elements and thus the occuring sums are in fact finite sums.

Any norm q on a vector space V produces a distance function (or metric) $d : V \times V \to \mathbb{R}$ via $d(x, y) = q(x - y)$. Thus V becomes a metric space and thus a topological space. The *open* sets $U \subset V$ are described by the property that to any point $x \in U$ there exists an ϵ-ball centered at x, $B_\epsilon(x) = \{y \in V \, ; \, q(y - x) < \epsilon\}$, contained entirely in U, i.e., $B_\epsilon(x) \subset U$ for some $\epsilon > 0$. The open sets so defined depend in general on the choice of norm on V. It is an important fact that for finite dimensional normed vector spaces all norms produce the same open sets. It can be easily checked that two norms q_k, $k = 1, 2$, on a vector space V give the same open sets iff there exist constants $c > 0$, $C > 0$ such that $cq_1(x) \leq q_2(x) \leq Cq_1(x)$ for all $x \in V$. This is clearly an equivalence relation. Thus we can rephrase that on a finite dimensional vector space all norms are equivalent. One should be aware that on a non-finite dimensional vector space two norms need not be equivalent: just take the norms q_1 and q_∞ in Example 1.4. Another important feature of finite dimensional normed vector spaces is that they are *complete*, i.e., every Cauchy sequence converges. Again, this is not the case in the non-finite dimensional case.

The usual concepts of convergency, limits, continuity etc. discussed in topology can now be formulated (and do not depend on the choice of norm in the finite dimensional situation). For example, the modified triangle inequality implies that any norm $q : V \to \mathbb{R}$ is a continuous function. From the multiplicative property of the operator norm one deduces that any linear map between finite dimensional normed vector spaces in continuous.

1.2. **The derivative.** In the sequel all vector spaces are assumed to be finite dimensional. We will denote the norm on any vector space V by the symbol $|x|$. We develop the basics of calculus on a normed vector space.

Definition 1.1. Let V, W be normed vector spaces, $U \subset V$ open and $f : U \to W$ a map (or function, whichever terminology one prefers). f is *differentiable* at $x_0 \in U$ if there exists a linear map $T : V \to W$ so that

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - T(x - x_0)}{|x - x_0|} = 0\,. \tag{6}$$

A useful equivalent formulation is the following

Lemma 1.1. *f is differentiable at x_0 if there exists a linear map $T : V \to W$ and a function $r : U \to W$ with $\lim_{x\to x_0} r(x) = 0$ such that*

$$f(x) = f(x_0) + T(x - x_0) + r(x)|x - x_0|$$

for $|x - x_0|$ sufficiently small, i.e., f can be approximated near x_0 by an affine (=inhomogeneous linear) map to more then first order.

From this two important facts follow immediately: firstly, if f is differentiable at x_0 then the linear map T is unique. Secondly, if f is differentiable at x_0 then the map f is continuous at x_0. To prove the first assertion we put $x = x_0 + tv$ for $v \in V$ and $t > 0$ small. Assuming there are two linear maps T_1, T_2 we get

$$T_1(v) - T_2(v) = (r_2(x) - r_1(x))|v|$$

for all v. Taking the limit $t \to 0$ we have $x \to x_0$ so that the right hand side becomes zero. For the second assertion we take the norm on both sides in the defining equation for differentiability to obtain

$$\begin{aligned} |f(x) - f(x_0)| &= |T(x - x_0) + r(x)|x - x_0|| \leq |T(x - x_0)| + |r(x)|x - x_0|| \\ &\leq (|T| + |r(x)|)|x - x_0| \end{aligned} \tag{7}$$

where we used properties of the norm and the operator norm. Since $\lim_{x\to x_0} r(x) = 0$ we obtain the usual ϵ, δ-characterization of continuity.

Note that even though the definitions (and calculations) involve a choosen norm, the property to be differentiable is independent of the norm (since all norms are equivalent). We introduce the following notation:

if f is differentiable at x_0 we denote the unique linear map T by $f'(x_0)$ and call it the *derivative* of f at x_0. Note that even though we use the same notation as for 1-variable calculus the derivative of a map in higher dimensions is a *linear* map $f'(x_0) : V \to W$. It is educational to compare this defintion to the one where the domain is an open intervall $I \subset \mathbb{R}$: let $f : I \to \mathbb{R}^n$ be a map (the image of such a map is a *curve* in $\mathbb{R}^n$) which is differentiable at some point $x_0 \in I$ in the above sense. Thus there exists a linear map $f'(x_0) : \mathbb{R} \to \mathbb{R}^n$ satisfying (6). Since a linear map is determined by its values on a basis and $\mathbb{R}$ has 1 as a canonical "basis vector" we can identify the linear map $f'(x_0)$ with the vector $f'(x_0)(1) \in \mathbb{R}^n$. Now it follows from (6) that

$$\lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)(1)\right) = 0\,,$$

which is exactly the definition of the derivative of a vector-valued function given in multivariable calculus (where the vector $f'(x_0)(1)$ is then interpreted as the "tangent vector" to the curve f in $\mathbb{R}^n$ at x_0). In the sequel we will *always* make this identification if the domain is an open subset of $\mathbb{R}$ without further mentioning.

1.3. **Directional derivatives and the Jacobi matrix.** When computing derivatives the notion of *directional derivatives* is useful.

Definition 1.2. Let $U \subset V$ be open, $f : U \to W$ a map and let $v \in V$. The *directional derivative* of f with respect to v at $x_0 \in U$ is defined by

$$D_v f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} \tag{8}$$

provided the limit exists.

If f is differentiable at x_0 then we conclude from (6) that directional derivatives of f with respect to any $v \in V$ exist and moreover

$$D_v f(x_0) = f'(x_0)(v)\,. \tag{9}$$

The converse is false as the example $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \frac{x^2 y}{x^2 + y^2}$, $f(0, 0) = 0$ shows: f has directional derivatives at (0,0) w.r.t. all $v \in \mathbb{R}^2$ but f is not differentiable at (0,0). Thus, as long as a map is differentiable at a point, we can compute its derivative using (9). In practical calculations one can choose bases $\mathbf{v} = (v_1, ..., v_n)$ in V and $\mathbf{w} = (w_1, ..., w_m)$ in W and express the linear map $f'(x_0)$ as a matrix w.r.t. that basis

$$f'(x_0)(\mathbf{v}) = \mathbf{w} J(f, x_0)\,, \tag{10}$$

or, using indices,

$$f'(x_0)(v_k) = \sum_{l=1}^{m} w_l J(f, x_0)^l_k\,, k = 1, ...n\,.$$

The matrix $J(f, x_0) \in \mathbf{M}(m, n, \mathbb{K})$ is called the *Jacobi matrix* of the map f at x_0. Its entry at the k-th column and l-th row

$$J(f, x_0)^l_k = D_{v_k} f^l(x_0)$$

is equal to the directional derivative w.r.t. v_k of the l-th component function $f^l : U \to \mathbb{K}$ of f w.r.t. the basis $\mathbf{w}$, i.e. $f = \sum_{l=1}^m w_l f^l$. In case $V = \mathbb{R}^n, W = \mathbb{R}^m$ one has the standard basis and then the directional derivatives w.r.t. the standard basis become the usual partial derivatives w.r.t the various "coordinate directions"

$$D_{e_k} f^l(x_0) = \frac{\partial f^l}{\partial x_k}(x_0) = \partial_k f^l(x_0)$$

so that

$$J(f, x_0)^l_k = \partial_k f^l(x_0)$$

is the matrix of all the partial derivatives, as in multivariable calculus.

1.4. **Higher order derivatives.** A map $f : U \to W$ of an open subset $U \subset V$ is called *differentiable* (on U) if f is differentiable at any point $x \in U$. Thus for each $x \in U$ we have the linear map $f'(x) \in Hom(V, W)$ and thus we get a map

$$f' : U \to Hom(V, W)$$

which we call the *derivative* or sometimes the *differential* of f. From Section 1.1 we know that $Hom(V, W)$ is also a normed vector space (e.g. in the operator norm) so that we can apply our theory to the map f'. If f' is a continuous map then we say that f is *continuously differentiable* on $U \subset V$ and write

$$C^1(U, W) := \{f : U \to W\,;\, f \,\text{continuously differentiable on}\, U\}\,.$$

The *second derivative* of f at $x_0 \in U$ is defined to be the derivative of f' at x_0, i.e.,

$$f''(x_0) := (f')'(x_0) \in Hom(V, Hom(V, W))$$

which now is a linear map from V to $Hom(V, W)$. From linear algebra we know that

$$Hom(V, Hom(V, W)) = Bil(V, W)\,,$$

where $Bil(V, W) = \{b : V \times V \to W\,;\, b\,\text{bilinear}\}$ is the vector space of bilinear maps. The isomorphism above is given by assigning a linear map $T \in Hom(V, Hom(V, W))$ the bilinear map $b(v_1, v_2) = T(v_1)(v_2)$. As before, if f' is differentiable at any $x \in U$ we say that f is *twice differentiable* (on U) and we obtain the second derivative map

$$f'' : U \to Hom(V, Hom(V, W)) = Bil(V, W)\,.$$

Again we have a map between normed vector spaces where one can check that the norm obtained on $Bil(V, W)$ under the isomorphism with $Hom(V, Hom(V, W))$ is given by $|b| = \sup_{v_1 \neq 0, v_2 \neq 0} \frac{|b(v_1, v_2)|}{|v_1||v_2|}$. If f'' is continuous we call f *twice continuously differentiable* and we write

$$C^2(U, W) := \{f : U \to W\,;\, f \,\text{twice continuously differentiable on}\, U\}\,.$$

The analog of the symmetry of mixed partial derivatives is the following

Lemma 1.2. *Let $f \in C^2(U, W)$. Then $f'' : U \to Sym^2(V, W)$ where we denote by $Sym^2(V, W) \subset Bil(V, W)$ the vector subspace of symmetric bilinear maps satisfying $b(v_1, v_2) = b(v_2, v_1)$.*

A fairly efficient way to compute second derivatives is the following: first one computes the derivative f'. For any $v_1 \in V$ the map

$$U \in x \to g_{v_1}(x) := f'(x)(v_1) \in W$$

is differentiable (assuming f is twice differentiable). Then the derivative of g_{v_1} yields the second derivative of f via

$$(g_{v_1})'(x)(v_2) = f''(x)(v_1, v_2)\,. \tag{11}$$

Continuing our boot-strapping we inductively define k-th order derivatives

$$f^{(k)} := (f^{(k-1)})' : U \to Hom(V, Hom(V, ..., Hom(V, W)) = Mult^k(V, W)$$

where $Mult^k(V, W) = \{b : V \times ... \times V \to W\,;\, b\, k\text{-multilinear}\}$ is the vector space of k-multilinear maps. Note that $Mult^1(V, W) = Hom(V, W)$ and $Mult^2(V, W) = Bil(V, W)$. We let

$$C^k(U, W) = \{f : U \to W\,;\, f \text{ k-times continuously differentiable on } U\,\}$$

and the above Lemma generalizes: if $f \in C^k(U, W)$ then $f^{(k)} \in Sym^k(V, W)$ where $Sym^k(V, W) \subset Mult^k(V, W)$ are those multilinear maps which are symmetric in all of its arguments, i.e., $b(v_{\sigma(1)}, v_{\sigma(2)}, ..., v_{\sigma(k)}) = b(v_1, ..., v_k)$ for all permutations $\sigma : \{1, ..., k\} \to \{1, ..., k\}$. Finally we let $C^\infty(U, W)$ be the intersection of all $C^k(U, W)$, i.e., maps which are differentiable of any order or infinitely often differentiable and shortly called *smooth* maps.

In multivariable calculus the notion of the *gradient* of a scalar valued function is somewhat important. This is to some extend misleading. To define the gradient one needs more structure on the vector space then just a norm, namely a non-degenerate symmetric bilinear form (inner product), not necessarily positive definite. Let $f : U \to \mathbb{K}$ be differentiable with derivative $f' : U \to Hom(V, \mathbb{K}) = V^*$ and let $<\,,\,>: V \times V \to \mathbb{K}$ be an inner product. Then we know from linear algebra that V and its dual V^* are isomorphic via $v \to < v, - >$. Using this isomorphism (which *depends on the choice of inner product*!) we define the gradient $\operatorname{grad}_x f \in V$ of f at $x \in U$ by

$$< \operatorname{grad}_x f, v >:= f'(x)(v)\,, v \in V\,.$$

Whereas the derivative f' is invariantly defined, the gradient function $\operatorname{grad} f : U \to V$ depends on the choice of inner product. When dealing with *Euclidean* vector spaces, i.e., $\mathbb{R}$-vector spaces with a positive inner product, the inner product (being positive definite) induces the norm $|x| = \sqrt{<x, x>}$ and there are no choices to make. On $\mathbb{R}^n$ one has a natural choice of inner product, the dot product, so that one often does not distinguish between $\operatorname{grad} f$ and f', at least not in calculus.

1.5. **Basic differentiation rules.** Without good formulas to differentiate products and compositions of maps the calculation of derivatives using the definition would be tedious. Like in calculus it follows from the limit properties that the derivative has the usual linearity properties

$$(\alpha f)' = \alpha f'\,, \qquad (f + g)' = f' + g'\,,$$

where $f, g : U \to W$, $U \subset V$ open, $\alpha \in \mathbb{K}$ and clearly $(\alpha f)(x) = \alpha f(x)$, $(f + g)(x) = f(x) + g(x)$.

Like in calculus we expect that we should be able to use the definition to calculate derivatives of constant maps, linear maps, bilinear maps and multilinear maps. Clearly, if $f : U \to W$ is a constant map, say $f(x) = w$ for all $x \in U$, then $f' = 0$ and thus all $f^{(k)} = 0$ for any $k \in \mathbb{N}$.

Derivative of a linear map. Let $f : U \to W$ be the restriction of a linear map $T \in Hom(V, W)$ to $U \subset V$. Then it follows from (6) that $f' = T$, i.e., $f'(x) = T \in Hom(V, W)$ for all $x \in U$. In particular, f' is a constant map and thus from the above consideration $f'' = 0$. Thus $f \in C^\infty(U, W)$ is infinitely often differentiable.

Derivative of a bilinear map. Let $b : U \times V \to W$ be a bilinear map between the normed vector spaces U, V, W. Then $b'(x, y) \in Hom(U \times V, W)$ is given by

$$b'(x, y)(u, v) = b(u, y) + b(x, v)\,.$$

This follows from the definition of the derivative (6). We see that the map $(x, y) \to b'(x, y)$ is a *linear* map (check!) between $U \times V$ and $Hom(U \times V, W)$. Thus we know from the above discussion that b'' must be constant. To explicitely calculate it we use (11) and obtain

$$b''(x, y)((u_1, v_1), (u_2, v_2)) = b(u_2, v_1) + b(u_1, v_2)$$

which obviously is independent of (x, y), i.e., constant. Thus the third and all higher derivatives vanish. A similar calculation for multilinear maps $b : V_1 \times ... \times V_n \to W$ gives

$$\begin{aligned} &b'(x_1, ..., x_n)(v_1, ...v_n) = \\ &\qquad = b(v_1, x_2, ..., x_n) + b(x_1, v_2, x_3, ..., x_n) + ... + b(x_1, ..., x_{n-1}, v_n)\,. \end{aligned}$$

With some patience one can also calculate all the higher order derivatives. We see that b' is an $n-1$-multilinear map if b has been n-multilinear. Thus inductively $b^{(n)}$ is constant and $b^{(n+1)} = 0$.

Chain rule. Let V_1, V_2, V_3 be normed vector spaces (always assumed finite dimensional) with $U_1 \subset V_1$ and $U_2 \subset V_2$ open subsets. Let $f : U_1 \to U_2$ be differentiable at $x \in U_1$ and let $g : U_2 \to V_2$ be differentiable at $f(x) \in U_2$. Then the composition $g \circ f : U_1 \to V_3$ is differentiable at $x \in U_1$ and the chain rule

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x) \tag{12}$$

holds. Note that the right hand side makes perfect sense as compositions of linear maps: $f'(x) \in Hom(V_1, V_2)$ and $g'(f(x)) \in Hom(V_2, V_3)$.

Product rule. Let V_1, V_2, V_3, W be normed vector spaces, $U \subset V_1$ open, $f : U \to V_2$, $g : U \to V_3$ maps and $b : V_2 \times V_3 \to W$ bilinear. If f and g are differentiable at $x_0 \in U$ then the map $b(f, g) : U \to W$ where $b(f, g)(x) := b(f(x), g(x))$, is differentiable at x_0 and

$$b(f, g)'(x_0)(v) = b(f'(x_0)(v), g(x_0)) + b(f(x_0), g'(x_0)(v)) \tag{13}$$

for all $v \in V_1$. To prove this formula we use the chain rule (12) together with the formula for the derivative of a bilinear map: note that $b(f, g)$ is the composition $b \circ (f, g)$ so that

$$\begin{aligned} (b \circ (f, g))'(x_0)(v) &= b'(f(x_0), g(x_0))(f'(x_0)(v), g'(x_0)(v)) \\ &= b(f'(x_0)(v), g(x_0)) + b(f(x_0), g'(x_0)(v))\,. \end{aligned}$$

1.6. **Inverse and implicit function theorems.** Let $U_1 \subset V_1$, $U_2 \subset V_2$ be open subsets of normed vector spaces. A map $f : U_1 \to U_2$ is said to be a C^k-*diffeomorphism*, $k \geq 1$, if

(i) $f : U_1 \to U_2$ is bijective and
(ii) f and $f^{-1} : U_2 \to U_1$ are k-times continuously differentiable.

Note that differentiability implies continuity so that diffeomorphisms are always homeomorphism and as such preserve all topological properties.

Since $f^{-1} \circ f = id_{U_1}$ and $f \circ f^{-1} = id_{U_2}$ we deduce from the chain rule

$$(f^{-1})'(f(x)) \circ f'(x) = id_{V_1} \, , x \in U_1$$

and

$$(f)'(f^{-1}(y) \circ (f^{-1})'(y) = id_{V_2} \, , y \in U_2 \, .$$

Thus the linear map $f'(x) \in Hom(V_1, V_2)$ is invertible with inverse $(f^{-1})'(f(x)) \in Hom(V_2, V_1)$. In particular, V_1 and V_2 must have the same dimension.

Theorem 1.3 (Inverse function theorem)**.** *Let V, W be normed vector spaces of the same dimension. Let $U \subset V$ be open and $f : U \to W$ be C^k and assume that $f'(x_0) \in Hom(V, W)$ is an invertible linear map at some $x_0 \in U$. Then there exist open sets $x_0 \in U_1 \subset U$ and $f(x_0) \in U_2 \subset W$ so that $f : U_1 \to U_2$ is a C^k-diffeomorphism.*

We say that a map $f : U \to W$, $U \subset W$ open, is a *local C^k-diffeomorphism near $x_0 \in U$* if there exists an open neighborhood $U_1 \subset U$ of x_0 and an open neighborhood $U_2 \subset W$ of $f(x_0)$ so that $f : U_1 \to U_2$ is a C^k-diffeomorphism. If $f : U \to W$ is a local C^k-diffeomorphism near *every point* in U we call f simply a *local C^k-diffeomorphism* (on U). From the above discussion and the inverse function theorem we deduce

Corollary 1.4. *A C^k map $f : U \to W$, $U \subset V$ open, is a local C^k-diffeomorphism near $x_0 \in U$ if and only if $f'(x_0) \in Hom(V_1, V_2)$ is invertible.*

Notice that a map $f : U \to W$ can be a *local* C^k-diffeomorphism (everywhere on U) without being a C^k-diffeomorphism as e.g. the map $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (e^x \cos y, e^x \sin y)$ shows. f has $f'(x, y)$ invertible for all (x, y) but f is neither injective nor surjective (see exercise).

A standard first application of the inverse function theorem is the *implicit function theorem*. First we need to introduce some notation. If $f : U_1 \times U_2 \to W$ is a differentiable map where, as usual, $U_1 \subset V_1$ and $U_2 \subset V_2$ are open subsets (of the normed vector spaces V_1, V_2) we define analogs of *partial derivatives* for f:

$$\partial_1 f(x, y) := (f \circ i_y)'(x) \in Hom(V_1, W) \, ,$$
$$\partial_2 f(x, y) := (f \circ i_x)'(y) \in Hom(V_2, W) \, ,$$

where $i_y : V_1 \to V_1 \times V_2$ is the inclusion map $i_y(x) = (x, y)$ and $i_x : V_2 \to V_1 \times V_2$ is the inclusion map $i_x(y) = (x, y)$. Like for the usual partial derivatives we see that $\partial_1 f(x, y)$ is nothing but the derivative of f in "x-direction keeping y constant" and correspondingly for $\partial_2 f(x, y)$. Since the derivative of $i_y(x) = (x, y)$ is $(i_y)'(x)(v) = (v, 0)$, which is the inclusion map $inc_1 : V_1 \to V_1 \times V_2$ (and correspondingly for i_x) we obtain from the chain rule

$$\partial_k f(x, y) = f'(x, y) \circ inc_k \in Hom(V_k, W) \, , k = 1, 2 \, .$$

This being said we can formulate

Theorem 1.5 (Implicit function theorem)**.** *Let V_1, V_2, W be normed vector spaces, $U_1 \subset V_1$, $U_2 \subset V_2$ open subsets and $f : U_1 \times U_2 \to W$ a C^k map. Assume that $\partial_2 f(x_0, y_0) \in Hom(V_2, W)$ is an invertible linear map for some $(x_0, y_0) \in U_1 \times U_2$. Then there exists an open neighborhood $B_1 \subset U_1$ of x_0 and a C^k map $g : B_1 \to U_2$*

with $g(x_0) = y_0$ satisfying $f(x, g(x)) = f(x_0, y_0) =: z_0$ for $x \in B_1$. Put more intuitively, the C^k equation $f(x, y) = z_0$ can be locally solved in a C^k manner explicitly for $y = g(x)$.

Of course, we all have seen this theorem used in calculus when one tries to solve an equation $f(x, y) = 0$ on $\mathbb{R}^2$ explicitly for $y = g(x)$ as a function of x. Geometrically this means that we want to write the curve $f(x, y) = 0$ locally as a graph $y = g(x)$ of some function g. The implict function theorem tells us that this can be done locally near points where the partial derivative $\frac{\partial f}{\partial y} \neq 0$. In general the theorem does not tell you how to find such a map g, it just guarantees its existence. In this light you should contemplate again the equation of the circle $f(x, y) = x^2 + y^2 - 1 = 0$.

The proof of the implicit function theorem is an easy application of the inverse function theorem: we introduce the map $F : U_1 \times U_2 \to V_1 \times W$ given by $F(x, y) = (x, f(x, y))$. Since $F'(x_0, y_0)(v_1, v_2) = (v_1, f'(x_0, y_0)(v_1, v_2)$ we deduce $F'(x_0, y_0)(v_1, v_2) = 0$ if and only if $v_1 = 0$ and $0 = f'(x_0, y_0)(0, v_2) = f'(x_0, y_0)(inc_2(v_2)) = \partial_2 f(x_0, y_0)(v_2)$. By assumtion $\partial_2 f(x_0, y_0)$ is invertible thus $v_2 = 0$. This shows that $F'(x_0, y_0) \in Hom(V_1 \times V_2, V_1 \times W)$ is injective. But $\dim V_2 = \dim W$, since $\partial_2 f(x_0, y_0) \in Hom(V_2, W)$ is invertible, thus $F'(x_0, y_0)$ is invertible. By the inverse function theorem F is a local diffeomorphism near (x_0, y_0). Hence there exist open neighborhoods $\tilde{U_1} \subset U_1$ of x_0, and $\tilde{U_2} \subset U_2$ of y_0, $B_1 \subset V_1$ of x_0 and $B_2 \subset W$ of z_0 so that $F : \tilde{U_1} \times \tilde{U_2} \to B_1 \times B_2$ is a C^k-diffeomorphism. Denote $F^{-1} = (h_1, h_2)$ and define $g(x) := h_1(x, z_0)$ which is a C^k-map $g : B_1 \to \tilde{U_1} \subset U_1$. Using $F \circ F^{-1} = id$ we get $f(x, g(x)) = z_0$.

Finally, under the assumptions of the implicit function theorem, we obtain the derivative of the map g by "implicit differentiation": applying the chain rule to the equation $f(x, g(x)) = z_0$ we get

$$\partial_1 f(x, g(x)) + \partial_2 f(x, g(x)) \circ g'(x) = 0$$

from which we deduce

$$g'(x) = -(\partial_2 f(x, g(x)))^{-1} \circ \partial_1 f(x, g(x)) .$$

You should check for yourself that the linear maps on the right hand side have the correct domains and ranges to be composed in the way they are.

1.7. **Normal forms of differentiable maps.** Recall that the *rank* of a linear map $T : V \to W$ between vector spaces of $\dim V = n, \dim W = m$ is defined to be the dimension of the vector subspace $T(V) = \{T(v) \,;\, v \in V\} \subset W$ given by the image of the linear map T. In particular, $\operatorname{rank} T \leq \min\{n, m\}$. If we chose bases in V and W and express T as a matrix $A \in \mathbf{M}(m, n, \mathbb{K})$ then $\operatorname{rank} T = r$ if and only if A has an $r \times r$ submatrix with non-zero determinant (this follows from the fact that $\operatorname{rank} T$ is also the dimension of the column space of A). Applying row and column operations to A one shows that T has rank r if and only if there exists bases in V, W so that the matrix A has the form $Ax = (x_1, ..., x_r, 0, ..., 0)^T \in \mathbb{K}^m$, $x \in \mathbb{K}^n$. Applying row and column operations is the same as applying invertible linear maps on the domain and range. Thus, we have characterized linear maps up to linear isomorphism of domain and range. A similar result, much in the same spirit, holds for differentiable maps.

Definition 1.3. We say a differentiable map $f : U \to W$ at $x_0 \in U$ has rank r at $x_0 \in U$ if its derivative $f'(x_0) \in Hom(V, W)$ at x_0 has $\operatorname{rank} f'(x_0) = r$. We will adopt the notation $\operatorname{rank}_{x_0} f := \operatorname{rank} f'(x_0)$. Note that $\operatorname{rank}_{x_0} f$ is the same as the rank of the Jacobi matrix $J(f, x_0)$ for any choice of bases.

Theorem 1.6 (Constant rank theorem)**.** *Let V, W be normed vector spaces of dimensions n and m, $U \subset V$ an open subset and $f : U \to W$ a C^k map. Assume that f has* constant rank r *on U, i.e.,* $\operatorname{rank}_x f = r$ *for all $x \in U$. To $x_0 \in U$ let $K := \ker f'(x_0) \subset V$ and $E := \operatorname{image} f'(x_0) \subset W$ be respectively the kernel and image of $f'(x_0)$. Then there exist open neighborhoods $B \subset U$ of x_0, $D \subset W$ of $f(x_0)$ and C^k-diffeomorphisms $\phi : B \to \tilde{B}$, $\psi : D \to \tilde{D}$, where $\tilde{B} \subset E \oplus K$ and $\tilde{D} \subset W$ are open, so that $\tilde{f} := \psi \circ f \circ \phi^{-1} : \tilde{B} \to \tilde{D} \subset W$ is of the form*

$$\tilde{f}(y, x) = y$$

for all $(y, x) \in \tilde{B} \subset E \oplus K$. Thus, up to diffeomorphisms of the domain and range, f looks very simple.

Before we get into the proof of this theorem let us discuss two important special cases, namely those where either $K = 0$ or $E = W$.

Definition 1.4. A C^k map $f : U \to W$, $U \subset V$ open, is called a C^k *immersion (submersion) at* $x_0 \in U$ if $f'(x_0) \in Hom(V, W)$ is injective (surjective). We call f a C^k-immersion (submersion) on U if f is a C^k-immersion (submersion) at all points in U.

Notice that $f'(x_0)$ is injective (surjective) if and only if its rank is maximal, i.e. $\operatorname{rank}_{x_0} f = \max(\dim V, \dim W)$, and $\dim V \leq \dim W$ ($\dim V \geq \dim W$) which is the same as $K = \ker f'(x_0) = 0$ ($E = \operatorname{image} f'(x_0) = W$). In case $\dim V = \dim W$ we have that $f'(x_0)$ is bijective, i.e., f is a local C^k-diffeomorphim near x_0. In this sense the notion of an immersion (submersion) generalizes the one of a local diffeomorphism.

Next we note that if f has maximal rank at a point $x_0 \in U$ then f has that same maximal rank in an open neighborhood $B \subset U$ of x_0. This follows from the fact that if the Jacobi matrix $J(f, x_0)$ has an $r \times r$ submatrix with non-zero determinant, by continuity of the determinant function and the continuity of the map $x \to J(f, x)$ (f is C^k, $k \geq 1$!), this same submatrix must have non-zero determinant in an open neighborhood of x_0, hence the rank cannot get smaller in this neighborhood (this is usually expressed by the phrase "the rank is lower semi-continuous"). Since we assumed the rank at x_0 to be maximal, it cannot get larger either, so it must stay constant in that neighborhood. Thus, if f is an immersion (submersion) at a point $x_0 \in U$ then f is an immersion (submersion) in some neighborhood of that point. We can thus apply the constant rank theorem 1.6 to immersions (submersions). We shall use the same notation as in that theorem.

Theorem 1.7 (Immersion theorem)**.** *Let $f : U \to W$ be a C^k-immersion at $x_0 \in U$ and let $E =$* $\operatorname{image} f'(x_0)$*. Then there exists open neighborhoods $B \subset U$ of x_0, $D \subset W$ of $f(x_0)$ and C^k-diffeomorphisms $\phi : B \to \tilde{B}$, $\psi : D \to \tilde{D}$, where $\tilde{B} \subset E$ and $\tilde{D} \subset W$ are open, so that $\tilde{f} := \psi \circ f \circ \phi^{-1} : \tilde{B} \to \tilde{D} \subset W$ is of the form*

$$\tilde{f}(y) = y$$

for all $y \in \tilde{B} \subset E$. Thus, up to C^k-diffeomorphisms of the domain and range, f is the restriction of the inclusion map $inc : E \to W$, $inc(y) = y$.

Theorem 1.8 (Submersion theorem)**.** *Let $f : U \to W$ be a C^k-submersion at $x_0 \in U$ and let $K = \ker f'(x_0)$. Then there exist open neighborhoods $B \subset U$ of x_0, $D \subset W$ of $f(x_0)$ and C^k-diffeomorphisms $\phi : B \to \tilde{B}$, $\psi : D \to \tilde{D}$, where $\tilde{B} \subset E \oplus K$ and $\tilde{D} \subset W$ are open, so that $\tilde{f} := \psi \circ f \circ \phi^{-1} : \tilde{B} \to \tilde{D} \subset W$ is of the form*

$$\tilde{f}(y, x) = y$$

for all $(y, x) \in \tilde{B} \subset E \oplus K$. Thus, up to diffeomorphisms of the domain and range, f is the restriction of the projection map $pr : E \oplus K \to E$, $pr(y, x) = y$.

Notice that both theorems contain the inverse function theorem 1.3 as the special case when $f'(x_0)$ is invertible. The proofs of the last two theorems are immediate applications of the constant rank theorem 1.6 and the discussion above: f being an immersion (submersion) at x_0 means that $\operatorname{rank}_{x_0} f = r$ is maximal and thus constant $\operatorname{rank}_x f = r$ for all x in an open neighborhood $B \subset U$ of x_0. We can now apply Theorem 1.6 keeping in mind that in the immersion case $K = 0$ and in the submersion case $E = W$.

From the submersion theorem we immediately deduce the following useful result whose proof is left as an exercise.

Corollary 1.9. *Let $U \subset V$ be open and $f : U \to W$ be a C^k-submersion. Then f is open, i.e., the image $f(B)$ of every open set $B \subset U$ is open in W.*

We conclude with the proof of the constant rank theorem. For notational ease we will ommit the attribute C^k. Let us choose direct summands $H \subset V$ and $F \subset W$ so that $V = H \oplus K$ and $W = E \oplus F$ and write $f = (f_1, f_2) : U \to E \oplus F = W$ and $x = (x_1, x_2) \in V = H \oplus K$. We first construct the domain diffeomorphism: consider the map $\phi : U \to E \oplus K$ defined by $\phi(x_1, x_2) = (f_1(x_1, x_2), x_2)$. Its derivative at x_0 is given by $\phi'(x_0)(v) = (f_1'(x_0)(v), v_2)$ for $v = (v_1, v_2) \in H \oplus K = V$. Whence $\phi'(x_0)(v) = 0$ if and only if $v_2 = 0$ and $f_1'(x_0)(v_1) = 0$. Since $f_1 = pr_E \circ f$, with $pr_E : W \to E$ the projection onto E along F, the chain rule gives us $f_1'(x_0) = pr_E \circ f'(x_0)$ where we used $pr_E'(x_0) = pr_E$ (compare to the section on the derivative of a linear map). Thus we have

$$f_1'(x_0)(v) = pr_E \circ f'(x_0)(v) = f'(x_0)(v)\,,$$

where the last equality is implied by the fact that $E = \operatorname{image} f'(x_0)$. Since any linear map is injective on any complementary subspace of its kernel we deduce from

$$0 = f_1'(x_0)(v_1) = f'(x_0)(v_1)\,, v_1 \in H$$

that $v_1 = 0$. Thus $\phi'(x_0) \in Hom(V, E \oplus K)$ is an injective linear map. On the other hand the dimension formula for linear maps $\dim \operatorname{image} f'(x_0) + \dim \ker f'(x_0) = \dim V$ gives $\dim V = \dim E \oplus K$. Thus $\phi'(x_0)$ is bijective and by the inverse function theorem 1.3 a local diffeomorphism near x_0, i.e., there exist open neighborhoods $B \subset U$ of x_0 and $\tilde{B} \subset E \oplus K$ of $\phi(x_0) = (f_1(x_0), (x_0)_2)$ so that $\phi : B \to \tilde{B}$ is a diffeomorphism. Now let $(y_1, x_2) = \phi(x_1, x_2) \in \tilde{B}$. Then

$$f \circ \phi^{-1}(y_1, x_2) = f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2)) = (y_1, h(y_1, x_2)) \tag{14}$$

where we used the defintion of $\phi(x_1, x_2) = (f_1(x_1, x_2), x_2)$ and put $h = f_2 \circ \phi^{-1} : \tilde{B} \to F$.

Our final task will be to construct a local diffeomorphism of the range W which will make $h = 0$. Notice that so far we have not used the fact that the rank of $f'(x)$ is *constant* on U. The derivative of $f \circ \phi^{-1}$ is given by

$$(f \circ \phi^{-1})'(z)(w, v) = (w, h'(z)(w, v)) = (w, \partial_1 h(z)(w) + \partial_2 h(z)(v)) \tag{15}$$

for $(w, v) \in E \oplus K$ and $z \in \tilde{B}$. From $(f \circ \phi^{-1})'(z) = f'(\phi^{-1}(z)) \circ (\phi^{-1})'(z)$, the assumtion that $f'(x)$ has constant rank $r = \dim E$ on $B \subset U$ and the fact that $(\phi^{-1})'(z)$ is a linear isomorphism we obtain that $(f \circ \phi^{-1})'(z)$ has constant rank r on $\tilde{B} \subset E \oplus K$. Thus $\dim \mathrm{image}(f \circ \phi^{-1})'(z) = r$. From (15) we have for each $z \in \tilde{B}$

$$\mathrm{image}(f \circ \phi^{-1})'(z) = \{(w, \partial_1 h(z)(w))\,;\, w \in E\} \oplus \{(0, \partial_2 h(z)(v))\,;\, v \in K\}\,.$$

Since $\dim\{(w, \partial_1 h(z)(w))\,;\, w \in E\} = r$ we deduce $\partial_2 h(z)(v) = 0$ for all $v \in K$ and $z \in \tilde{B}$. Note that by shrinking $\tilde{B}$ we may assume that $\tilde{B} = \tilde{B}_1 \times \tilde{B}_2 \subset E \oplus K$ with $\tilde{B}_1 \subset E$ and $\tilde{B}_2 \subset K$ open and connected (in fact, we could choose them to be ϵ-neighborhoods). For $z = (z_1, z_2) \in \tilde{B}$ we have $\partial_2 h(z)(v) = (h \circ inc_2)'(z_2)(v)$, where $inc_2 : \tilde{B}_2 \to \tilde{B}$ is the inclusion map, and thus $\partial_2 h(z)(v) = 0$ implies $(h \circ inc_2)(z_2) = c$ constant on $\tilde{B}_2$. Hence $h(z_1, z_2) = h(z_1)$ does not depend on $z_2 \in \tilde{B}_2$ and we define the map $\psi : \tilde{B}_1 \times F \to W$ by $\psi(y_1, y_2) = (y_1, y_2 - h(y_1))$. From $\psi'(y)(w) = (w_1, w_2 + h'(y_1)(w_1))$ we deduce that ψ is a local diffeomorphism on the open subset $\tilde{B}_1 \times F \subset W$ which contains the point $f(x_0)$ since $\tilde{B}_1$ contains the point $f_1(x_0)$. Thus there exist open neighborhoods $D \subset \tilde{B}_1 \times F$ of $f(x_0)$ and $\tilde{D} \subset W$ of $\psi(f(x_0))$ on which $\psi : D \to \tilde{D}$ is a diffeomorphism. Note that continuity of f implies that $f^{-1}(D) \cap B$ is an open neighborhood of $f(x_0)$. Hence we may assume that $f(B) \subset D$ (by replacing B with $f^{-1}(D) \cap B$). It remains to check that $\psi \circ f \circ \phi^{-1}$, the compositions now being well-defined, has the desired property. If $(z_1, z_2) \in \tilde{B} \subset E \oplus K$ then

$$\psi \circ f \circ \phi^{-1}(z_1, z_2) = \psi(z_1, h(z_1)) = (z_1, h(z_1) - h(z_1)) = (z_1, 0) = z_1 \in E \subset W\,,$$

where we used (14) and the definition of ϕ.

2. Submanifolds of normed vector spaces

Submanifolds are the natural generalizations of curves and surfaces in 3-space which we encounter in multivariable calculus. They also will provide the link to abstract manifolds.

2.1. **Submanifolds described by local equations.** In what follows the letters V, W will denote finite dimensional normed vector spaces.

Definition 2.1. Let $M \subset V$ be a subset of V. A point $p_0 \in M$ is said to be a C^k*-smooth point of dimension* m in M if there exists an open neighborhood $U \subset V$ of p_0 in V and a C^k-map $g : U \to \tilde{V}$, $\dim \tilde{V} = \dim V - m$, which is a submersion at p_0 so that

$$M \cap U = g^{-1}(0) = \{p \in U\,;\, g(p) = 0\}\,.$$

Since a submersion at p_0 is always a submersion on an open neighborhood of p_0 we might equally well have defined a smooth point by requiring the existence of a C^k-submersion $g : U \to \tilde{V}$ with $M \cap U = g^{-1}(0)$. Sometimes we will refer to g as a *defining function* for M.

The subset $M \subset V$ is called a C^k*-submanifold* of dimension m of V if all points $p \in M$ are C^k-smooth points of dimension m. We write $\dim M = m$ and call $\operatorname{codim} M := \dim V - m$ the *codimension* of M in V.

If we do not want to emphazise the degree of differentiablility $k \geq 1$ and in the case of $k = \infty$ we will simply say "smooth point" and "submanifold". Let us look at some examples:

Example 2.1. Every point $x \in V$ is a zero dimensional submanifold: its defining function is $g(x) = x - x_0$.

Example 2.2. On the other extreme every open subset $M \subset V$ is a submanifold of codimension zero: the defining function is the zero map on M with range $\mathbb{R}^0 = \{0\}$.

Example 2.3. Let $g : \mathbb{R}^2 \to \mathbb{R}$ be the map $g(x_1, x_2) = x_1^2 - x_2$. Then $g'(x)(v_1, v_2) = 2x_1v_1 - v_2$, i.e., $g'(x) = (2x_1, -1)$ w.r.t. the standard basis. Hence $\operatorname{rank} g'(x) = 1$ everywhere on $\mathbb{R}^2$ and thus g is a submersion on $\mathbb{R}^2$. By the above definition $M = g^{-1}(0) = \{x \in \mathbb{R}^2\,;\, x_2 = x_1^2\}$ is a 1-dimensional submanifold of $\mathbb{R}^2$.

Perhaps slightly more interesting is the following

Example 2.4. Let $<,>: V \times V \to \mathbb{R}$ be a non-degenerate symmetric bilinear form and consider $M = \{x \in V\,;\, < x, x >= 1\}$. We will show that M is a submanifold of codimension 1 in V. A defining function for M is $g(x) =< x, x > -1$ which has derivative $g'(x)(v) = 2 < x, v >$ for $x, v \in V$. Since $<,>$ is non-degenerate for $x \neq 0$ there always is some $v \neq 0$ with $< x, v >\neq 0$. Thus $g'(x) \neq 0$ for $x \neq 0$,i.e., $\operatorname{rank} g'(x) = 1$ for $x \neq 0$ which is to say that $g : V \setminus \{0\} \to \mathbb{R}$ is a submersion. Since $M \subset V \setminus \{0\}$ we see that M is a submanifold of codimension 1.

It might be constructive to remind ourselves that we can always chose a basis $\mathbf{v} = (v_1, ..., v_n)$ in V so that the matrix $(< v_i, v_j >) = \operatorname{diag}(1, ..., 1, -1, ..., -1) =: I_{s,n-s}$ has s many $+1$'s and $n - s$ many -1's in the diagonal. On $V = \mathbb{R}^n$ with the standard basis the inner product given by $I_{s,n-s}$ is usually called a *Lorentzian* inner product of *signature* $(s, n-s)$, $s \in \{0, ..., n\}$ and M is referred to as a Lorentz sphere. Of course, the case $s = 0$ gives the standard Euclidean inner product and M is just the standard sphere. As an exercise you should calculate and draw the various Lorentz spheres in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Example 2.5. If $M \subset V$ is a submanifold and $O \subset M$ is an open set (in the subspace topology) then $O \subset V$ is a submanifold with $\dim O = \dim M$. This follows at once from the definition and the fact that $O = M \cap U$ with $U \subset V$ open.

In explicit examples we often have an obvious defining function for our canditate submanifold M. It may happen that this function is not a submersion but only has constant rank. This turns out to be sufficient for M to still be a submanifold.

Theorem 2.1. *Let $g : V \to W$ be a map and denote $M = g^{-1}(0)$. Assume that there exists an open neighborhood $U \subset V$ containing $M \subset U$ on which $g : U \to W$ is a C^k-map of constant rank. Then M is a submanifold of* $\dim M = \dim V - \operatorname{rank} g$.

Note that this theorem contains two special cases: if the map g is a submersion at points $p \in M$ then g is a submersion in an open neighborhood of V near p. The union of those neighborhoods over M give an open set $U \supset M$ on which g is a submersion and thus of constant rank, namely $\operatorname{rank} g = \dim W$. We then recover what we already knew that is, M is a submanifold of $\dim M = \dim V - \dim W$. On the other hand, if g is an immersion at points of M then (by the same reasoning) there is an open neighborhood $U \supset M$ on which g is an immersion and thus of constant rank $\dim V$. Since g is locally injective $g^{-1}(0)$ consists of isolated points which, according to the first example, is a submanifold of dimension zero (agreeing with the dimension formula of the theorem).

Proof. We have to show that every point $p_0 \in M$ is a smooth point of dimension $\dim V - \operatorname{rank} g$. Let $E = \operatorname{image} g'(p_0)$ and $K = \ker g'(p_0)$. Then the constant rank theorem 1.6 gives open neighborhoods $B \subset U$ of p_0 and $D \subset W$ of $g(p_0) = 0$ and diffeomorphisms $\phi : B \to \tilde{B}$, $\psi : D \to \tilde{D}$ where $\tilde{B} \subset E \oplus K$ and $\tilde{D} \subset W$ are open, so that $\tilde{g} = \psi \circ g \circ \phi^{-1}$ is of the form $\tilde{g}(z_1, z_2) = (z_1, 0) \in E \oplus K$. Consider the C^k-map $h : B \to E$ given by $h(p) = \tilde{g}_1(\phi(p))$. Since $\operatorname{rank} g = \dim E$ we also have $\operatorname{rank} \tilde{g} = \dim E$ (they differ by diffeomorphisms) and thus $\operatorname{rank} h = \dim E$. The latter says that $h : B \to E$ is a submersion. Moreover $h^{-1}(0) = (\psi \circ g)^{-1}(0) = (g_{|B})^{-1}(0) = B \cap M$ which says that h is a defining function for M near p_0 and whence p_0 is a smooth point of dimension $\dim V - \operatorname{rank} g$. $\square$

Another useful observation is that the inverse image of a submanifold under a submersion is again a submanifold (this is generalization of the fact that the inverse image of a point under a submersion is a submanifold, which follows from the definition of a submanifold). We leave the proof as an exercise and just state the result:

Theorem 2.2. *Let $g : V \to W$ be a C^k-map and let $N \subset W$ be a submanifold. If g is a submersion at all points in $M := g^{-1}(N) \subset V$ then M is a submanifold of V. The dimension of M is* $\dim M = \dim V - \dim W + \dim N = \dim V/W + \dim N$ *and its codimension is* $\operatorname{codim} M = \operatorname{codim} N$.

2.2. **Submanifolds described by local parametrizations.** An important feature of submanifolds is that they can locally be "parametrized". This will allow us to extend the notions of differentiability to functions defined on submanifolds and also motivate the notion of an "abstract manifold". We already know from calculus that a curve in the plane can either be given implicitly as the zero set of a function or it can be given as a parametrized curve. Our submanifolds so far are defined "implicitly" by defining functions. We now show that we could have also defined them in terms of "local parametrizations".

Theorem 2.3. *Let $M \subset V$ and $p_0 \in M$. Then p_0 is a smooth point of dimension m if and only if there exists an injective immersion $f : D \to V$, where $D \subset W$ is an open subset of a normed vector space W of dimension m, such that $f(D) \subset M$ is an open neighborhood of $p_0 \in M$ (in the subspace topolgy of $M \subset V$). We call such f a* local parametrization *of M near p_0. By translating in W we may always assume $f(0) = p_0$.*

Parametrization Theorem. If $p_0 \in M$ is a smooth point of dimension m then there exists an open neighborhood $U \subset V$ of p_0 and a submersion $g : U \to \tilde{V}$, $\dim \tilde{V} = \dim V - m$, with $g^{-1}(0) = M \cap U$. By the submersion theorem 1.8 there are open neighborhoods $B \subset U$ of p_0, $D \subset \tilde{V}$ of 0 and diffeomorphisms $\phi : B \to \tilde{B}$, $\psi : D \to \tilde{D}$, with $\tilde{B} =\subset K \times \tilde{V}$ and $\tilde{D} \subset \tilde{V}$ open neighborhoods of 0 (we may assume, after possible translations, that $\phi(p_0) = 0$ and $\psi(0) = 0$) so that $\tilde{g} = \psi \circ g \circ \phi^{-1} : \tilde{B} \to \tilde{D}$ has $\tilde{g}(x, y) = y$. As usual we have put $K = \ker g'(p_0) \subset V$. We can also assume that $\tilde{B} = \tilde{B_1} \times \tilde{B_2}$ with $\tilde{B_1} \subset K$ and $\tilde{B_2} subset W$ open neigborhoods of 0. Now define $f : \tilde{B_1} \to V$ by $f(x) = \phi^{-1}(x, 0)$. Clearly, f is an injective immersion and

$$f(\tilde{B_1}) = \psi^{-1}(\tilde{B_1} \times \{0\}) = \psi^{-1}(\tilde{g}^{-1}(0)) = (g_{|B})^{-1}(\psi^{-1}(0)) = (g_{|B})^{-1}(0) = M \cap B$$

which says that $f(\tilde{B_1})$ is open in the subspace topology of M. Since $\dim K + \operatorname{rank} g'(p_0) = \dim K + \dim \tilde{V} = \dim K + \dim V - m = \dim V$ we have $\dim K = m$ so that f is defined on an open neighborhood of 0 of the m-dimensional normed vector space K.

To prove the converse let $f : D \to V$, $D \subset W$ open, be an injective immersion with $f(D) = U \cap M$, $U \subset V$ open, and $f(0) = p_0$. We need to show that $p_0 \in M$ is a smooth point of $\dim W$. By the immersion theorem 1.7 there are open neighborhoods $0 \in O \subset D$, $p_0 \in B \subset U$ and diffeomorphisms $\phi : O \to \tilde{O}$, $\psi : B \to \tilde{B}$ where $\tilde{O} \subset E = \operatorname{image} f'(0) \subset V$, $\tilde{B} \subset V$ are open, so that $f(O) \subset B$ and $\tilde{f} = \psi \circ f \circ \phi^{-1}$ is the restriction of the inclusion map $inc : E \to E \times F = V$ to $\tilde{O} \subset E$. Here $F \subset V$ is a direct summand of E in V and $\tilde{B} = \tilde{B}_1 \times \tilde{B}_2 \subset E \times F$ which we may always assume. Now let $\tilde{B}' = \tilde{O} \times \tilde{B}_2$ and put $B' = \psi^{-1}(\tilde{B}')$. Then $f(O) = B' \cap M$ since f is injective. Define $g : B' \to \tilde{B}_2 \subset F$ by $g = pr_2 \circ \psi$ where $pr_2 : \tilde{O} \times \tilde{B}_2 \to \tilde{B}_2$. Then g is a submersion on B', since pr_2 is one and ψ is a diffeomorphism. Moreover, $p_0 \in B'$ and

$$g^{-1}(0) = \psi^{-1}(\tilde{O} \times \{0\}) = f(\phi^{-1}(\tilde{O})) = f(O) = B' \cap M\,.$$

Finally $\tilde{B}_2 \subset F$ and $\dim F = \dim V - \dim E = \dim V - \dim W$ so that $p_0 \in M$ is a smooth point of dimension $\dim W$. □

As an immediate Corollary we obtain

Corollary 2.4. *$M \subset V$ is an m-dimensional submanifold if and only if near each point $p \in M$ there exists a local parametrization f whose domain is an open subset of some* $\dim V - m$*-dimensional normed vector space.*

In principle we now could develop calculus on submanifolds. As a first step towards this we need the notion of differentiability of a function on a submanifold $M \subset V$. Of course, this notion should be compatible with the one we already have on open sets (which are submanifolds as well) and we expect that restrictions of differentiable functions on V to M should be differentiable.

Definition 2.2. Let $M \subset V$ be a C^k-submanifold and $h : M \to \tilde{V}$ a map into some normed vector space. We say h is *k-times differentiable*, $k \geq 1$, or briefly C^k, if for all local parametrizations $f : D \to M$ of M we have that $h \circ f : D \to \tilde{V}$ is a C^k-map.

In many applications the maps we encounter are only defined on an open subset $O \subset M$. Those are submanifolds (of the same dimension then M) and thus we can apply the above definition. This comes down to checking differentiablilty of $h \circ f$ for those local parametrizations f of M whose images are contained in O. Note that any local parametrization $f : D \to M$ of M whose image intersects $O \subset M$ gives rise to a local parametrization (again called) $f : D \cap f^{-1}(O) \to O$ of O. We only have to see that $D \cap f^{-1}(O)$ is open in $D \subset W$: but $O = M \cap U$ with $U \subset V$ open and since $f(D) \subset M$ we have $f^{-1}(O) = f^{-1}(U)$ which is open in W.

In order not to have to check differentiability for *all* local parametrizations we need

Lemma 2.5. *Let $M \subset V$ be a C^k submanifold and let $f_i : D_i \to M$ be two local parametrizations with $D_i \subset W_i$. Then*

$$f_2^{-1} \circ f_1 : f_1^{-1}(f_1(D_1) \cap f_2(D_2)) \to f_2^{-1}(f_1(D_1) \cap f_2(D_2))$$

is a C^k-diffeomorphism.

The important consequence of this lemma is

Corollary 2.6. *Let $M \subset V$ be a C^k submanifold, $h : M \to \tilde{V}$ a map and let $f_i : D_i \to M$ be two local parametrizations with $D_i \subset W_i$. Then $h \circ f_1$ is C^k if and only if $h \circ f_2$ is C^k. Thus it suffices to check differentiability of h only w.r.t.* some *family of local parametrizations whose images cover M.*

The corollary follows immediately from lemma 2.5: $h \circ f_1 = (h \circ f_2) \circ (f_2^{-1} \circ f_1)$. We now come to the proof of lemma 2.5:

Proof. Since $f_i(D_i) \subset M$ are open we get by the arguments above on restrictions of parametrizations to open subsets of M that $f_i^{-1}(f_1(D_1) \cap f_2(D_2))$ are open in $D_i \subset W_i$. Shrinking D_i to $f_i^{-1}(f_1(D_1) \cap f_2(D_2))$ we may assume $f_1(D_1) = f_2(D_2) = U \cap M$. Since both f_i are injective they are bijective onto their images and thus $f_2^{-1} \circ f_1$ is bijective with inverse $f_1^{-1} \circ f_2$. It suffices to show that $f_2^{-1} \circ f_1$ is C^k since switching the roles of f_1 and f_2 will also show that $f_1^{-1} \circ f_2$ is C^k. To $x_0 \in D_1$ the immersion theorem 1.7 applied to f_1 gives open subsets $D \subset D_1$, $B \subset U$ and diffeomorphisms $\phi : D \to \tilde{D} \subset E$, $\psi : B \to \tilde{B} \subset V$, where as usual $E = \operatorname{image} f_1'(0)$, so that $\psi \circ f_1 \circ \phi^{-1} = inc : \tilde{D} \to \tilde{D} \times \tilde{B}_2$. Here we assume as before that $\tilde{B} = \tilde{D} \times \tilde{B}_2$ with $\tilde{B}_2 \subset F$ open and F a direct summand of E in V. Then $f_1(D) = B \cap M$. Now

$$f_2^{-1} \circ f_1 = f_2^{-1} \circ (\psi^{-1} \circ inc \circ \phi)\,,$$

but the map $\psi \circ f_2 : D' \to \tilde{B}$, where $D' = f_2^{-1}(B) \subset D_2$ is open, takes values in $\tilde{D} \times \{0\}$ and as such is a local C^k-diffeomorphism. Thus its inverse $(\psi \circ f_2)^{-1} : \tilde{D} \to B'$ is C^k and we can rewrite

$$f_2^{-1} \circ f_1 = (\psi \circ f_2)^{-1} \circ \phi$$

which is C^k as the composition of C^k maps. $\square$

An obvious way to get differentiable maps on a submanifold $M \subset V$ is by restricting differentiable maps of V (or of an open neighborhood $U \supset M$ in V) to M: if $h : U \to \tilde{V}$ is C^k and $f : D \to M \subset U \subset V$ is a local parametrization then

$$(h_{|M}) \circ f = h \circ f : D \to \tilde{V}$$

is C^k as the composition of C^k-maps. It is left as an exercise that locally any C^k map on M is obtained this way which in particular implies that a C^k-map $h : M \to W$ is continuous (in the subspace topology of M).

If $f : D \to M \subset V$ is a local parametrization then $O := f(D) \subset M$ is open and $f^{-1} : O \to D \subset W$ is C^k since $f^{-1} \circ f = id_{|D}$ is C^k (here we used f as the parametrization w.r.t. which we check differentiability of f^{-1}). We call the inverse map $f^{-1} : O \to D \subset W$ of a local parametrization $f : D \to M$ a *local chart* or *local coordinate* for M and O the *chart* or *coordinate domain*. Since $f^{-1} : O \to D \subset W$ is continuous f (and f^{-1}) is a homeomorphism.

If $M \subset V$ is an open subset then it has the (global) parametrization $id_{|M} : M \to M$. Thus a map $h : M \to \tilde{V}$ is C^k (definition 2.2) if and only if $h \circ id_{|M} = h : M \to V$ is C^k. Thus the notion of differentiablility on submanifolds coincides for open subsets with the usual notion of C^k.

From the above discussions we see that an m-dimensional C^k-submanifold $M \subset V$ admits an open covering $\{O_i\,;\, i \in I\}$ by chart domains with chart maps $\phi_i : O_i \to D_i$. The $D_i \subset W_i$ are open subsets of m-dimensional normed vector spaces W_i and the ϕ_i are homeomorphisms. The *chart transition functions* or *coordinate transition functions*

$$\phi_i \circ \phi_j^{-1} : \phi_j(O_i \cap O_j) \to \phi_i(O_i \cap O_j)$$

between the open subsets $\phi_j(O_i \cap O_j) \subset W_j$ and $\phi_i(O_i \cap O_j) \subset W_i$ are C^k-diffeomorphisms. As you can see, none of the properties listed refers to the surrounding space V (except the notion of M being a submanifold). In the next chapter we will use exactly those properties to define the notion of an *abstract* C^k-manifold.

Example 2.6. A standard way to parametrize the sphere $S^n(r) = \{x \in \mathbb{R}^{n+1}\,;\, |x|^2 = r^2\}$ of radius $r > 0$ in $\mathbb{R}^{n+1}$ is by the inverse stereographic projection maps. Recall that by example 2.4 $S^n(r) \in \mathbb{R}^{n+1}$ is a submanifold of dimension n (take $<,>$ to be the standard dot-product on $\mathbb{R}^{n+1}$ scaled by $1/r^2$). We define two maps $f_\pm : \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

$$f_\pm(x) = (\frac{2xr^2}{|x|^2 + r^2}, \pm\frac{r(|x|^2 - r^2)}{|x|^2 + r^2})\,.$$

You should check that $f_\pm$ are injective immersions with images image $f_\pm = S^n(r) \setminus \{\pm re_{n+1}\}$ which are both open in $S^n(r)$. The corresponding chart maps $\phi_\pm$ are the stereographic projections from the "north pole" re_{n+1} and the "south pole" $-re_{n+1}$ onto the equatorial hyperplane $\mathbb{R}^n \times \{0\}$. You should compute the coordinate transition function.

A map $h : M_1 \to M_2$ between two C^k-submanifolds $M_i \subset V_i$ is C^k if it is a C^k-map regarded as a map $h : M_1 \to V_2$. We expect that composing h by any local chart map ϕ of M_2 the resulting map $\phi \circ h$ is C^k:

Lemma 2.7. *Let $h : M_1 \to M_2$ be a map between two C^k-submanifolds $M_i \subset V_i$. Then h is C^k as a map $h : M_1 \to V_2$ if and only if for any local parametrization*

$f : D \to M_2 \subset V_2$, $D \subset W_2$ open, the map $f^{-1} \circ h : h^{-1}(f(D)) \to D \subset W_2$ is C^k (in the sense of definition 2.2).

Proof. Obviously, if $f^{-1} \circ h : h^{-1}(f(D)) \to D \subset W_2$ is C^k then $f \circ f^{-1} \circ h = h$ is C^k as a map into V_2 since f is. The converse can be most easily seen by recalling that $f^{-1} : f(D) \to D \subset W_2$ is locally the restriction of a C^k-map ϕ defined in some open neighborhood $U \subset V_2$ of a point in $f(D)$ to $f(D) \cap U$. □

2.3. **The tangent bundle of a submanifold.** Even though we now understand the notion of a differentiable function on a submanifold $M \subset V$ there is still a lot missing for an effective calculus on submanifolds. The obvious first thing to look at is the analog of the derivative which will be the last concept we are going to introduce on submanifolds per se. It leads naturally to the notion of the *tangent bundle*. This will give us an explicit example on which we can base our intuition when developing the abstract manifold setup.

Definition 2.3. Let $M \subset V$ be a subset. To $p \in M$ the subset

$$T_pM := \{v \in V\,; \text{there exists } \epsilon > 0 \text{ and a } C^k\text{-curve}\, \gamma : (-\epsilon, \epsilon) \to M \\ \text{with } \gamma(0) = p \text{ and } \gamma'(0) = v\}$$

is called the *C^k-tangent space* (or simply the *tangent space*) of M at $p \in M$. Elements $v \in T_pM$ are called *tangent vectors* of M at p.

Unless $p \in M$ is a smooth point of M we do not expect this set to have much structure.

Theorem 2.8. *Let $M \subset V$ and assume that $p_0 \in M$ is a C^k-smooth point of dimension m. Then:*

(i) *$T_{p_0}M \subset V$ is a vector subspace of dimension m.*
(ii) *If $g : U \to \tilde{V}$ is a defining function for M near p_0 then $T_{p_0}M = \ker g'(0)$.*
(iii) *If $f : D \to M$ is a local parametrization of M near p_0 with $f(x_0) = p_0$ then $T_{p_0}M = \operatorname{image} f'(x_0)$.*

The affine space $p_0 + T_{p_0}M \subset V$ is what we usually draw in pictures as the "geometric" tangent space.

Example 2.7. Consider the sphere w.r.t. a (not necessarily positive definit) inner product $<,>: V \times V \to \mathbb{R}$ given by $M = \{x \in V\,;\, < x, x >= 1\}$ (compare example 2.4). We have one defining function $g(x) =< x, x > -1$ which has $g'(x)(v) = 2 < x, v >$ and so

$$T_xM = \ker g'(x) = \{v \in V\,;\, < x, v >= 0\}\,.$$

In the case $<,>$ is the usual dot product in $V = \mathbb{R}^n$ this says that the tangent space at a point x of the sphere consits of all the vectors perpendicular to the "radius" vector x.

Proof. (i) follows from either (ii) or (iii). Let us start with the latter: first we show that $T_{p_0}M \subset \operatorname{image} f'(x_0)$. Let $v = \gamma'(0)$ for some C^k-curve $\gamma : (-\epsilon, \epsilon) \to M$. Since γ is continuous we may assume that $\operatorname{image}\gamma \subset f(D)$. From lemma 2.7 $c = f^{-1} \circ \gamma : (-\epsilon, \epsilon) \to D$ is a C^k-curve in D with $c(0) = x_0$. Putting $X = c'(0)$ we obtain

$$f'(x_0)(X) = (f \circ c)'(0) = \gamma'(0) = v$$

so that $v \in \operatorname{image} f'(x_0)$. To show that $\operatorname{image} f'(x_0) \subset T_{p_0}M$ we take $\gamma(t) = f(x_0 + tX)$, then $\gamma : (-\epsilon, \epsilon) \to M$ is a C^k-curve in M with $\gamma(0) = p_0$ and $\gamma'(0) = f'(x_0)(X)$.

To verify (ii) let $\gamma : (-\epsilon, \epsilon) \to M$ be a C^k-curve in M with $\operatorname{image}\gamma \subset U$, $\gamma(0) = p_0$ and $\gamma'(0) = v$. Then $g \circ \gamma = 0$ and differentiating this we get $g'(x_0)(v) = 0$. Thus $T_{p_0}M \subset \ker g'(p_0)$. From (iii) we know that $T_{p_0}M = \operatorname{image} f'(x_0)$ which is a $\operatorname{rank} f'(x_0) = m$ dimensional vector subspace of V. Since $\dim \ker g'(p_0) = m$ we conclude $T_{p_0}M = \ker g'(p_0)$. □

Definition 2.4. Let $M \subset V$ be a C^k-submanifold. Then

$$TM := \{(p, v) \in M \times V \,;\, v \in T_pM\} \subset V \times V$$

is called the *tangent bundle* of M.

Example 2.8. Let us look at the case where $M \subset V$ is an open subset. Then $T_pM = V$ for all $p \in M$ and thus $TM = M \times V$.

Theorem 2.9. *Let $M \subset V$ be a C^k-submanifold of $\dim M = m$ wit $k \geq 2$. Then $TM \subset V \times V$ is a C^{k-1}-submanifold of $\dim TM = 2m$.*

Proof. We show that every point $(p, v) \in TM$ is a C^{k-1}-smooth point of dimension $2m$ by constructing local parametrizations from local parametrizations $f : D \to M$, $D \subset W$ open, of M. The idea is that locally M "looks like" the open set $D \subset V$ and thus TM "looks like" $D \times W$. Let us carry this idea out and define

$$\Phi : D \times W \to TM \qquad \Phi(x, \xi) = (f(x), f'(x)(\xi))\,.$$

Then Φ is well defined since we know from theorem 2.8 that $f'(x)(\xi) \in T_{f(x)}M$. We have to show that Φ is an injective immersion whose image $\Phi(D \times W) \subset TM$ is open in the subspace topology $TM \subset V \times V$. If $\Phi(x, \xi) = \Phi(y, \eta)$ then $f(x) = f(y)$ and $f'(x)(\xi) = f'(y)(\eta)$. Since f is injective the former gives $x = y$. But f is also an immersion, i.e. $f'(x)$ is injective, so that the latter implies $\xi = \eta$. Next we show that Φ is an immersion by computing

$$\Phi'(x, \xi)(\eta_1, \eta_2) = (f'(x)(\eta_1), f''(x)(\xi, \eta_1) + f'(x)(\eta_2)$$

where $(x, \xi) \in D \times W$ and $(\eta_1, \eta_2) \in W \times W$. From this expression together with the fact that $f'(x)$ is injective we immediately get $\ker \Phi'(x) = 0$. Finally

$$\operatorname{image} \Phi = \bigcup_{x \in D} \{f(x)\} \times T_{f(x)}M = \bigcup_{p \in f(D)} \{p\} \times T_pM = (U \times V) \cap TM$$

where $f(D) = U \cap M$ with $U \subset V$ open. This shows that $\operatorname{image}\Phi$ is an open subset of TM. Clearly, Φ is C^{k-1} if f was C^k and $\dim TM = 2 \dim W = 2 \dim M$. □

The local parametrizations of TM constructed above are sometimes called *bundle parametrizations* and their inverses *bundle charts*. You should calculate their transition functions as an exercise.

The tangent bundle carries a natural projection map $\pi : TM \to M$, $\pi(p, v) = p$, called the *tangent bundle projection*, which is simply the restriction of the projection $pr_1 : V \times V \to V$ and thus differentiable. The inverse images, or *fibers*, $\pi^{-1}(p) = T_pM$ are the individual tangent spaces of M. Later on we shall investigate auxiliary manifolds over a base manifold admitting a "projection map" whose fibers are vector spaces in more detail (see chapter on vector bundles).

With the tagent bundle at hand we can define the derivative of a map between two submanifolds:

Definition 2.5. Let $M_i \subset V_i$ be two C^k-submanifolds and let $h : M_1 \to M_2$ be a C^k-map. The *derivative* (or *differential*) of h is defined by

$$h' : TM_1 \to TM_2 \qquad h'(p, v) = (h(p), (h \circ \gamma)'(0))$$

where $\gamma : (-\epsilon, \epsilon) \to M$ is a C^k-curve with $\gamma(0) = p$ and $\gamma'(0) = v \in T_pM$. That this is actually well defined, i.e., not dependent on the curve γ chosen, will be part of the theorem below. It will be convenient to denote the second component of h' by $h'_p(v) = (h \circ \gamma)'(0)$.

Theorem 2.10. *Let $h : M_1 \to M_2$ be a C^k-map between two C^k-submanifolds $M_i \subset V_i$ and denote the tangent bundle projections by $\pi_i : TM_i \to M_i$. Then the derivative $h' : TM_1 \to TM_2$ is well defined and C^{k-1}. Furthermore, $\pi_2 \circ h' = h \circ \pi_1$, i.e., $h'_p(v) \in T_{h(p)}M_2$ for $v \in T_pM_1$ and for all $p \in M_1$ the map $h'_p : T_pM_1 \to T_{h(p)}M_2$ is linear.*

Proof. To see that h' is well defined extend h to a C^k-map $\tilde{h}$ defined in some open neighborhood of $p \in M_1$ in V_1. Then

$$h'_p(v) = (h \circ \gamma)'(0) = (\tilde{h} \circ \gamma)'(0) = \tilde{h}'(p)(v)$$

which only depends on v and not on the curve γ chosen with $\gamma'(0) = v$.

In order to verify that h' is C^{k-1} we take a local parametrization $\Phi : D \times W \to TM_1$ of the form $\Phi(x, \xi) = (f(x), f'(x)(\xi))$ where $f : D \to M_1$, $D \subset W$ open, is a local parametrization of M_1. Then one easily checks that

$$(h' \circ \Phi)(x, \xi) = ((h \circ f)(x), (h \circ f)'(x)(\xi))$$

so that the map $(h' \circ \Phi) : D \times W \to TM_2 \subset V_2 \times V_2$ is C^{k-1}.

To check that $h'_p(v) = (h \circ \gamma)'(0) \in T_{h(p)}M_2$ we note that $h \circ \gamma$ is a C^k-curve in M_2 with $(h \circ \gamma)(0) = h(p)$. Thus, by definition of the tangent space, $h'_p(v) \in T_{h(p)}M_2$.

Finally we show that $h'_p : T_pM_1 \to T_{h(p)}M_2$ is linear. Take $v_1, v_2 \in T_pM_1$ and $a, b \in \mathbb{R}$. Represent v_i by curves γ_i in M_1, i.e., $\gamma_i(0) = p$ and $\gamma'_i(0) = v_i$, and $av_1 + bv_2$ by a curve γ. Extend as before h to a C^k-map $\tilde{h}$ defined in some open neighborhood of $p \in M_1$ in V_1. Then

$$\begin{aligned} h'_p(av_1 + bv_2) =& (h \circ \gamma)'(0) = (\tilde{h} \circ \gamma)'(0) = \tilde{h}'(p)(av_1 + bv_2) = \\ & a\tilde{h}'(p)(v_1) + b\tilde{h}'(p)(v_2) = a(\tilde{h} \circ \gamma_1)'(0) + b(\tilde{h} \circ \gamma_2)'(0) = \\ & a(h \circ \gamma_1)'(0) + b(h \circ \gamma_2)'(0) = ah'_p(v_1) + bh'_p(v_2) \end{aligned}$$

which finishes the proof. $\square$

Using the tangent bundle and the derivative h' of a map between submanifolds the chain rule becomes more natural to state:

Theorem 2.11. *Let $M_i \subset V_i$, $i = 1, 2, 3$, be C^k-submanifolds and let $h_1 : M_1 \to M_2$ and $h_2 : M_2 \to M_3$ be C^k-maps. Then $h_2 \circ h_1 : M_1 \to M_3$ is C^k and $(h_2 \circ h_1)' : TM_1 \to TM_3$ is given by the "chain rule"*

$$(h_2 \circ h_1)' = h'_2 \circ h'_1 \,.$$

Proof. You should convince *yourself* hat the composition of C^k-maps on submanifolds is again C^k. To proof the chain rule let $p \in M_1$ and $\gamma : (-\epsilon, \epsilon) \to M_1$ be a C^k-curve with $\gamma(0) = p$ and $\gamma'(0) = v$. Then $h_1 \circ \gamma$ is a C^k-curve in M_2 representing the tangent vector $(h_1 \circ \gamma)'(0)$. Thus

$$\begin{aligned}(h_2' \circ h_1')(p, v) =& h_2'(h_1(p), (h_1 \circ \gamma)'(0)) = \\ & (h_2(h_1(p)), (h_2 \circ (h_1 \circ \gamma))'(0)) = \\ & ((h_2 \circ h_1)(p), ((h_2 \circ h_1) \circ \gamma)'(0)) = \\ & (h_2 \circ h_1)'(p, v)\,.\end{aligned}$$

□

In case when $M_i \subset V_i$ are open subsets we already know that their tangent bundles $TM_i = M_i \times V_i$. The derivative $h_1' : TM_1 \to TM_2$ is given by $h_1'(p, v) = (h_1(p), (h_1 \circ \gamma)'(0))$ where we can choose $\gamma(t) = p + tv$ for small t. Thus

$$h_1'(p, v) = (h_1(p), h_1'(p)(v))$$

and the chain rule above gives

$$\begin{aligned}(h_2 \circ h_1)'(p, v) =& h_2'(h_1(p), h_1'(p)(v)) = (h_2(h_1(p)), h_2'(h_1(p))(h_1'(p)(v))) = \\ & (h_2(h_1(p)), (h_2'(h_1(p)) \circ h_1'(p))(v))\end{aligned}$$

which recovers in the second component (the first component is redundant) the usual chain rule

$$(h_2 \circ h_1)'(p)(v) = (h_2'(h_1(p)) \circ h_1'(p))(v)\,.$$

After having developed submanifolds and their basic calculus with relativly little effort we could now start to study differential-topological and geometrical questions. The aim of this course though is to aquaint you with the more general notion of an abstract manifold which does have advantages. As an example let us look at real projective space

$$\mathbb{R}P^n = \text{all lines through the origin in } \mathbb{R}^{n+1} = S^n/(\pm 1)$$

where the last quotient space is given by identifying antipodal points, x and $-x$, on the sphere $S^n \subset \mathbb{R}^{n+1}$ (a line through the origin cuts the sphere in an antipodal pair of points). From topology we know that $\mathbb{R}P^n$ can be given the quotient topology in which the natural projection $\pi : S^n \to \mathbb{R}P^n$, $\pi(x) = \{x, -x\}$, becomes continuous. With our methods so far it is not clear how to make $\mathbb{R}P^n$ into a submanifold (even though the sphere is a very simple submanifold of $\mathbb{R}^{n+1}$ there seems to be no way to see $\mathbb{R}P^n$ as a submanifold of $\mathbb{R}^{n+1}$). One way to get a submanifold which is homeomorphic to $\mathbb{R}P^n$ is the following (the details are left as an exercise): consider the map $f : S^n \to V$, $f(x) = xx^T$, into $V = \text{symmetric}\,(n+1) \times (n+1)$-matrices, where we view $x \in \mathbb{R}^{n+1}$ always as column vectors and x^T is the transpose, i.e., a row vector. The map f is differentiable, $f(x) = f(y)$ if and only if $y = -x$, and the image $M := f(S^n) \subset V$ is a submanifold of dimension n. Thus $f : \mathbb{R}P^n \to M$ is differentiable and bijective and furthermore, its inverse is continuous also. This exhibits $\mathbb{R}P^n$ as a submanifold in the space of symmetric matrices. Of course, when studying $\mathbb{R}P^n$ one does not necessarily want to work with this model of $\mathbb{R}P^n$. This prompts the notion of an "intrinsic" differentiable model of $\mathbb{R}P^n$. How to achieve this will be discussed in the next chapter.

3. Abstract manifolds

To define an abstract manifold we will simply use the properties of chart maps for a submanifold as discussed in the end of section 2.2 as axioms. It might be helpful to have those in mind as we proceed.

3.1. **Differentiable structures.**

Definition 3.1. Let M be a set. A *chart* (or *local coordinate*) on M is a pair (U, ϕ) where $U \subset M$ is a subset of M, $\phi : U \to W$ is an injective map into some (finite dimensional) normed vector space W and $\phi(U) \subset W$ is an *open* subset. In particular, $\phi : U \to \phi(U)$ is bijective. We call U the *chart (or coordinate) domain* and ϕ the *chart (or coordinate) map*. Note that the normed vector space W comes with the chart even though we do not explicitely mention it (to avoid overbearing notation). If $p \in M$ lies in the chart domain U we say (U, ϕ) is a *chart near* p.

Two charts (U_i, ϕ_i) are C^k*-compatible* if $\phi_i(U_1 \cap U_2) \subset W_i$ are open and $\phi_2 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2)$ is a C^k-diffeomorphism. This implies that also $\phi_1 \circ \phi_2^{-1}$ is a C^k-diffeomorphims so that compatibility of charts is a symmetric notion. Charts whose chart domains are disjoint, $U_1 \cap U_2 = \emptyset$, are always compatible. We allow $k = 0, 1, 2, ..., \infty, \omega$ where C^0-diffeomorphisms are homeomorphisms and C^ω stands for real analytic. Note that if two charts (U_i, ϕ_i) are C^k*-compatible* the vector spaces W_i must have the same dimension: this follows from the inverse function theorem 1.3 in case $k \neq 0$ and the invariance of domain theorem for $k = 0$.

A C^k*-atlas* of M is given by a collection of pairwise C^k-compatible charts $\mathcal{A} = \{(U_i, \phi_i)\,;\, i \in I\}$ whose chart domains cover M, i.e., $\bigcup_{i \in I} U_i = M$. Two C^k-atlases $\mathcal{A}_i$ are *equivalent* if their union $\mathcal{A}_1 \cup \mathcal{A}_2$ is again a C^k-atlas, i.e., if every chart in $\mathcal{A}_1$ is C^k-compatible with every chart in $\mathcal{A}_2$. We shall see below that this is indeed an equivalence relation. The equivalence classes $\mathcal{D}$ are called C^k*-differentiable* structures on M.

Given a C^k-differentiable structure $\mathcal{D}$ on M any C^k-atlas $\mathcal{A} \in \mathcal{D}$ is said to be an atlas *representing* $\mathcal{D}$, in other words, $\mathcal{A}$ represents $\mathcal{D}$ if $\mathcal{D} = [\mathcal{A}]$. The union $\mathcal{A}_{\max} = \bigcup_{\mathcal{A} \in \mathcal{D}} \mathcal{A}$ of all atlases in a C^k-differentiable structure is called the *maximal atlas* representing $\mathcal{D}$. By construction, $\mathcal{A}_{\max}$ is the largest C^k-atlas representing $\mathcal{D}$, i.e., any C^k-atlas $\mathcal{A}$ representing $\mathcal{D}$ is contained in $\mathcal{A}_{\max}$.

The concept of a differentiable structure is just a formal way to express the perhaps more intuitive idea of adding compatible charts to a given atlas untill one reaches a maximal atlas.

Definition 3.2. A C^k*-manifold* is a set M together with a C^k-differentiable structure $\mathcal{D}$.

In examples the differentiable structure $\mathcal{D}$ is always given by a representing atlas $\mathcal{A}$ so that $\mathcal{D} = [\mathcal{A}]$. The statement " let (U, ϕ) (or $\phi : U \to W$) be a chart of M" will always implicitely mean that the chart belongs to a representing atlas of the differentiable structure on M.

Definition 3.3. Let M be a C^k-manifold with differentiable structure $\mathcal{D}$. The *dimension at* $p \in M$ is defined by

$$\dim_p M := \dim W$$

for any chart $\phi : U \to W$ (belonging to the differentiable structure).

Notice that $\dim_p M$ is well-defined independently of the chart chosen, since the vector spaces W of compatible charts all have the same dimension (see above).

Example 3.1. Any submanifold of a normed vector space is a manifold by the above definition where the differentiable structure is represented by the atlas coming from local parametrizations (compare end of section 2.2). This already gives us a wealther of non-trivial examples of manifolds.

Next we discuss real projective space $\mathbb{R}P^n$ for which there was no obvious way to see it as a submanifold.

Example 3.2. Recall that $\mathbb{R}P^n$ is the set of lines through the origin in $\mathbb{R}^{n+1}$, i.e., a point $p \in \mathbb{R}P^n$ is given by a line $p = \mathbb{R}x$ with $0 \neq x \in \mathbb{R}^{n+1}$. Two vectors $x, y \in \mathbb{R}^{n+1}$ give the same line if and only if $y = ax$ for some $0 \neq a \in \mathbb{R}$. This leads to the description of $\mathbb{R}P^n$ as $(\mathbb{R}^{n+1} \setminus \{0\})/\mathbb{R}_*$ where points in projective space are equivalence classes $[x] \in (\mathbb{R}^{n+1} \setminus \{0\})/\mathbb{R}_*$. Note that representatives $y \in [x]$ are only determined up to multiplications by non-zero $a \in \mathbb{R}_*$. It is customary to call any $y = (y_0, ..., y_n) \in [x]$ *homogeneous coordinates* of the point $[x]$. Geometrically you should think of $\mathbb{R}P^n$ as $\mathbb{R}^n$ compactified by adding on "points at infinity" which can be represented by the set of all directions (i.e., lines) in $\mathbb{R}^n$, thus $\mathbb{R}P^n = \mathbb{R}^n \cup \mathbb{R}P^{n-1}$. For instance, the projective line $\mathbb{R}P^1$ is just $\mathbb{R}$ added on one point at infinity so that we can think of $\mathbb{R}P^1$ as a circle. The projective plane $\mathbb{R}P^2$ is $\mathbb{R}^2$ glued on an $\mathbb{R}P^1$, i.e., a circle, at infinity. How this circle is glued to $\mathbb{R}^2$ is something you might want to contemplate.

We define a C^ω-differentiable structure on $\mathbb{R}P^n$ by giving an atlas consisting of $n+1$ many charts: for $i = 0, ..., n$ let the chart domains be

$$U_i = \{[x] \in \mathbb{R}P^n \, ; \, x_i \neq 0\} \subset \mathbb{R}P^n$$

and define as chart maps

$$\phi_i : U_i \to \mathbb{R}^n \, , \qquad \phi([x]) = (x_0/x_i, ..., x_{i-1}/x_i, x_{i+1}/x_i, ..., x_n/x_0) \, .$$

Since ϕ_i is bijective with inverse $\phi_i^{-1} : \mathbb{R}^n \to U_i$ given by

$$\phi_i^{-1}(t_1, ..., t_n) = [t_1, ..., t_i, 1, t_{i+1}, ..., t_n]$$

(U_i, ϕ_i), $i = 0, \ldots, n$, are indeed charts. To show that they form an atlas we check their compatibility: let $0 \leq j < k \leq n$ then

$$\phi_j(U_j \cap U_k) = \{t \in \mathbb{R}^n \, ; \, t_k \neq 0\}$$
$$\phi_k(U_j \cap U_k) = \{t \in \mathbb{R}^n \, ; \, t_{j+1} \neq 0\}$$

which are both open subsets of $\mathbb{R}^n$. The coordinate transition map $\phi_k \circ \phi_j^{-1}$: $\phi_j(U_j \cap U_k) \to \phi_k(U_j \cap U_k)$ is given by

$$\begin{aligned}(\phi_k \circ \phi_j^{-1})(t) &= \phi_k([t_1, ..., t_j, 1, t_{j+1}, ..., t_n]) = \\ &= (t_1/t_k, ..., t_j/t_k, 1/t_k, t_{j+1}/t_k, ..., t_{k-1}/t_k, t_{k+1}/t_k, ..., t_n/t_k)\end{aligned}$$

which is a C^ω-diffeomorphism. Thus $\mathcal{A} = \{(U_i, \phi_i) \, ; \, i = 0, ..., n\}$ is a C^ω-atlas which makes $\mathbb{R}P^n$ into a C^ω-manifold. The charts given are frequently called *affine charts* for $\mathbb{R}P^n$ the reason being that each U_i presents an "ordinary" $\mathbb{R}^n$ inside $\mathbb{R}P^n$ via the map ϕ_i^{-1}. We will see later that this differentiable structure is "natural" in many ways.

Before we go on to discuss the topology of a manifold let us proof that the notion of equivalence of atlases is indeed an equivalent relation: reflexivity and symmetry are obvious from the defintion. If we have three atlases $\mathcal{A}_i$ so that $\mathcal{A}_1$ is equivalent to $\mathcal{A}_2$ and $\mathcal{A}_2$ is equivalent to $\mathcal{A}_3$ we need to see that $\mathcal{A}_1$ is equivalent to $\mathcal{A}_3$. Let $(U_i, \phi_i) \in \mathcal{A}_i$, $i = 1, 3$, be two charts. If the chart domains do not intersect then they are compatible charts. If $U_1 \cap U_3 \neq \emptyset$ then we have to show that $\phi_i(U_1 \cap U_3) \subset W_i$, $i = 1, 3$, are open and $\phi_3 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_3) \to \phi_3(U_1 \cap U_3)$ is a C^k-diffeomorphism. Openess can be shown as follows: any point $x_0 \in \phi_i(U_1 \cap U_3)$ is of the form $\phi_i(p_0)$ for $p_0 \in U_1 \cap U_3$. Take a chart $(U_2, \phi_2) \in \mathcal{A}_2$ with $p_0 \in U_2$. Since (U_2, ϕ_2) is compatible with both charts (U_i, ϕ_i), $i = 1, 3$, the subsets $\phi_2(U_2 \cap U_i) \in W_2$ and $\phi_i(U_2 \cap U_i) \in W_i$ are open and $\phi_i \circ \phi_2^{-1} : \phi_2(U_2 \cap U_i) \to \phi_i(U_2 \cap U_i)$, $i = 1, 3$, are C^k-diffeomorphisms. But the subset

$$\phi_2(U_1 \cap U_2 \cap U_3) = \phi_2(U_2 \cap U_1) \cap \phi_2(U_2 \cap U_3)$$

which is open in W_2. Applying the C^k-diffeomorphism $\phi_i \circ \phi_2^{-1}$ we obtain the open neighborhood $\phi_i(U_1 \cap U_2 \cap U_3) \subset \phi_i(U_1 \cap U_3) \subset W_i$ of x_0. To see that $\phi_3 \circ \phi_1^{-1}$ is a C^k-diffeomorphism it suffices to show this locally (since the map is already a bijection). But on the open subset $\phi_1(U_1 \cap U_2 \cap U_3) \subset W_1$ we can express

$$\phi_3 \circ \phi_1^{-1} = (\phi_3 \circ \phi_2^{-1}) \circ (\phi_2 \circ \phi_1^{-1}),$$

which is the composition of C^k-diffeomorphism, and as such a C^k-diffeomorphism.

3.2. **Manifold topology.** Manifolds come with a natural topology given by their differentiable structure.

Definition 3.4. Let M be a C^k-manifold with differentiable structure $\mathcal{D}$. We call a subset $O \subset M$ *open* if $\phi(O \cap U) \subset W$ is open for all charts $(U, \phi) \in \mathcal{A}$ for any atlas $\mathcal{A}$ representing the differentiable structure $\mathcal{D} = [\mathcal{A}]$.

As stated this definition, though only dependent on the differentiable structure, is not particularily useful to work with. We would have to test openess against all charts from all equivalent atlases. In practice, a manifold ususally is given by a specific atlas. That we only have to test against a particular atlas is the content of the next

Lemma 3.1. *Let M be a C^k-manifold with differentiable structure $\mathcal{D}$ and let $\mathcal{A} \in \mathcal{D}$. A subset $O \subset M$ is open if and only if $\phi(O \cap U) \subset W$ is open for all charts $(U, \phi) \in \mathcal{A}$.*

Proof. Let $(U_1, \phi_1) \in \mathcal{A}_1$ be a chart in an equivalent atlas $\mathcal{A}_1 \in \mathcal{D}$. We have to show that $\phi_1(O \cap U_1) \subset W_1$ is open. Let $x_0 \in \phi_1(O \cap U_1)$ (if $\phi_1(O \cap U_1) = \emptyset$ it is open and we are done) and let $p_0 = \phi_1^{-1}(x_0) \in O \cap U_1$. Since $\mathcal{A}$ is an atlas there exists a chart $(U, \phi) \in \mathcal{A}$ with $p_0 \in U$. The subset

$$\phi(O \cap U \cap U_1) = \phi(U_1 \cap U) \cap \phi(O \cap U) \subset W$$

is open since $\phi(U_1 \cap U) \subset W$ is open by definition of compatible charts and $\phi(O \cap U) \subset W$ is open by assumption. Moreover, $\phi_1 \circ \phi^{-1} : \phi(U \cap U_1) \to \phi_1(U \cap U_1)$ is a C^k-diffeomorphism so that its image $\phi_1(O \cap U \cap U_1) \subset \phi_1(O \cap U_1)$ of the open subset $\phi(O \cap U \cap U_1) \subset \phi(U \cap U_1)$ is open (in W_1) and contains x_0. Thus x_0 is an inner point of $\phi_1(O \cap U_1) \subset W_1$. Since x_0 was chosen arbitrarily we conclude that $\phi_1(O \cap U_1) \subset W_1$ is open. $\square$

We leave it as an exercise to verify that the collection of all open subsets of M indeed form a topology. We call this toplogy $\mathcal{T}_M$ the *manifold topology* of M.

Given a chart (U, ϕ) for M the chart domain U is open by definition and the chart map $\phi : U \to \phi(U) \subset W$ is a homeomorphism: since any open subset $O \subset U$ is also open in M and thus $\phi(O) \subset W$ is open we see that ϕ is an open map. To check that ϕ is continuous we take an open subset $B \subset \phi(U)$ and have to verify that $O := \phi^{-1}(B)$ is open in M. Let (U_1, ϕ_1) be a compatible chart then $U_1 \cap O = U_1 \cap U \cap O$ (note that $O \subset U$) and thus $\phi(U_1 \cap O) = \phi(U_1 \cap U) \cap B$, which is open in W. Moreover, $\phi_1 \circ \phi^{-1}$ is a C^k-diffeomorphism so that its image $\phi_1(O \cap U_1)$ of $\phi(U_1 \cap O)$ is an open subset of W_1.

From this we immediately conclude that the manifold topology is "locally euclidean", i.e., each point on M has an open neighborhood which is homeomorphic to an open neighborhood in some $\mathbb{R}^n$ (recall that any n-dimensional vector space is linearly homeomorphic to $\mathbb{R}^n$). Thus $\mathcal{T}_M$ has a countable neighborhood base (first countability axiom), is locally compact (each point has a compact neighborhood) and is locally path connected (so that the connected components of M are path connected). In general the manifold topology is neither Hausdorff nor second countable (i.e., there is no countable base for the topology).

Recall that the dimension of a C^k-manifold M at a point $p \in M$ (definition 3.3) is given by $\dim_p M = \dim W$ where $\phi : U \to W$ is a chart near $p \in M$. Thus, $\dim_p M = \dim_q M$ for all $p, q \in U$. Since chart domains are open this says that the map $p \mapsto \dim_p M$ is locally constant on M and thus constant on connected components.

Definition 3.5. Let M be a connected C^k-manifold. The *dimension* of M is given by $\dim M := \dim_p M$ for $p \in M$.

We have seen above that chart maps are homeomorphism in the manifold topology. This property already characterizes the manifold topology which is a very useful fact when identifying an already known topology as the manifold topology.

Lemma 3.2. *Let M be a C^k-manifold with atlas $\mathcal{A}$. A topology $\mathcal{T}$ on M is the manifold topology if and only if for all $(U, \phi) \in \mathcal{A}$ the chart domain $U \in \mathcal{T}$ and the chart map $\phi : U \to \phi(U) \subset W$ is a homeomorphism (where, of course, U carries the induced topology from $\mathcal{T}$).*

Proof. Let $\mathcal{T}$ be a topology for which the chart maps are homeomorphisms. We need to show that $\mathcal{T} = \mathcal{T}_M$. If $O \in \mathcal{T}_M$ then $O \cap U \in \mathcal{T}_M$ and hence $\phi(O \cap U) \subset W$ is open for all charts $(U, \phi) \in \mathcal{A}$ (since ϕ is a homeomorphism in the manifold topology). Since ϕ is also a homeomorphism w.r.t. the topology $\mathcal{T}$ we conclude that $O \cap U = \phi^{-1}(\phi(O \cap U)) \in \mathcal{T}$. Thus $O = \bigcup_{(U,\phi) \in \mathcal{A}} O \cap U$ is open w.r.t. $\mathcal{T}$. The other inclusion is shown similarily. □

Let us conclude this paragraph with some examples:

Example 3.3. If $M \subset V$ is a C^k-submanifold then the manifold topology is the subspace topology induced from V: this follows immediately from the previous lemma and the fact that the differentiable structure on M is given by the inverses of parametrizations which we have shown in section 2.2 to be homeomorphisms w.r.t. the subspace topology on $M \subset V$. Thus submanifolds of (finite dimensional) normed vector spaces are always Hausdorff and second countable.

Example 3.4. An example of a non-Hausdorff manifold is the set M of vertical lines in the punctured plane $\mathbb{R}^2 \setminus \{0\}$. For each non-zero x-value there is exactly one such line, but there are two (half) lines at $x = 0$ so that $M = \mathbb{R} \cup \{0'\}$. The two charts $(\mathbb{R}, id)$ and $(\mathbb{R}_* \cup \{0'\}, \phi)$, where $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$ and $\phi_{|\mathbb{R}_*} = id_{|\mathbb{R}_*}$, $\phi(0') = 0$, are easily checked to be C^ω-compatible. Thus M becomes a C^ω-manifold. Since the open neighborhoods of $0 \in M$ are just the usual open neighborhoods of $0 \in \mathbb{R}$ and the open neighborhoods of $0' \in M$ are punctured open neighborhoods $U \setminus \{0\}$ of $0 \in \mathbb{R}$ union $0'$ we conclude that 0 and $0'$ cannot be seperated, i.e., M is not Hausdorff. Regarding M as the set of equivalence classes of points in $\mathbb{R}^2 \setminus \{0\}$ where $x \sim y$ if and only if $x_1 = y_1 \neq 0$ or $\operatorname{sign}(x_2) = \operatorname{sign}(y_2)$ for $x_1 = y_1 = 0$ you should check, using the previous lemma, that the manifold topology on M is the quotient topology on $\mathbb{R}^2 \setminus \{0\}/\sim$. In particular, M is connected (images of connected sets under continuous maps are connected) and of dimension 1.

Example 3.5. Real projective space $\mathbb{R}P^n$ can be viewed as the set of equivalence classes $(\mathbb{R}^{n+1} \setminus \{0\})/\mathbb{R}_*$ and as such has the quotient topology $\tilde{\mathcal{T}}$ as a natural topology. We will see that this is in fact the manifold topology. In particular, $\mathbb{R}P^n$ is a connected, second countable, Hausdorff (assuming you have shown this in topology, otherwise do it now) manifold of dimension n. Recall that the quotient topology is the finest topology on $\mathbb{R}P^n$ making the coset projection $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}P^n$, $\pi(x) = [x]$, continuous, i.e., a subset $O \subset \mathbb{R}P^n$ is open w.r.t. $\tilde{\mathcal{T}}$ if and only if $\pi^{-1}(O) \subset \mathbb{R}^{n+1} \setminus \{0\}$ is open. Let $\mathcal{A} = \{(U_i, \phi_i)\,;\, i = 0, ..., n\}$ be the atlas of affine charts described in example 3.2. Then $\pi^{-1}(U_i) = \{x \in \mathbb{R}^{n+1} \setminus \{0\}\,;\, x_i \neq 0\}$ which is an open subset of $\mathbb{R}^{n+1} \setminus \{0\}$. Thus our chart domains are open in the quotient topology. Furthermore, ϕ_i is continuous in the quotient topology if and only if $\phi_i \circ \pi : \pi^{-1}(U_i) \to \mathbb{R}^n$ is continuous. But the latter map sends $(x_0, ..., x_n)$ to $(x_0/x_i, ..., x_{i-1}/x_i, x_{i+1}/x_i, ..., x_n/x_i)$ which is obviously continuous. Since $\phi_i^{-1} = \pi \circ f_i$ where $f_i : \mathbb{R}^n \to \mathbb{R}^{n+1} \setminus \{0\}$ is the continuous map $f_i(t) = (t_1, ..., t_i, 1, t_{i+1}, ..., t_n)$ also ϕ_i^{-1} is continuous. Thus $\phi_i : U_i \to \mathbb{R}^n$ is a homeomorphism w.r.t. the quotient toplogy on $\mathbb{R}P^n$ so that by lemma 3.2 the manifold topology of $\mathbb{R}P^n$ is the quotient topology. Finally note that π restricts to a continuous (surjective) map $\pi : S^n \to \mathbb{R}P^n$, where $S^n \subset \mathbb{R}^{n+1} \setminus \{0\}$ carries the subspace topology, so that $\mathbb{R}P^n$ is in fact compact.