MATH 437 - Manifolds

ARessegetes Stery

January 25, 2024

Contents

1   Preliminary Results from Calculus

This section is dedicated to proving the following two theorems, which has various implications in differential geometry:

Theorem 1.1 (Inverse Function Theorem). *Let X, Y be finite-dimensional normed vector space, $\Omega \subseteq X$ an open set, and $f : \Omega \to Y$ a continuously differentiable function. Let $x_0 \in \Omega$. If $Df|_{x_0}$ is invertible, then f is locally C^1-invertible at x_0 (i.e. f^{-1} is C^1-continuous), whose derivative satisfies*

$$Df^{-1}\big|_{f(x)} = \left(Df|_x\right)^{-1}$$

Theorem 1.2 (Implicit Function Theorem). *Let $\Omega \subseteq \mathbb{R}^n \times R^m$ be open, with $F \in C^1\Omega, \mathbb{R}^m$ s.t. $F(x_0, y_0) = 0$ for $(x_0, y_0) \in \Omega$. If*

$$\det DF(x_0, -)\big|_{y_0} \neq 0$$

then there exists $\varepsilon \in \mathbb{R}$, and $g \in C^1(B_\varepsilon(x_0), \mathbb{R}^m)$ s.t.

$$F(x, g(x)) = 0 \qquad \forall\, x \in B_\varepsilon(x_0)$$

The general setup is to first prove the inverse function theorem, and use it to vastly simplify the proof for the implicit function theorem. An alternative approach is to first prove the implicit function theorem, and then express the derivative of the inverse function in an implicit manner.

First setup some lemmas for further use in the proof:

Lemma 1.3 (Contraction Mapping Principle). *Let X be a complete normed vector space, and M be a closed subset of X. If $f : M \to M$ is Lipschitz continuous with Lipschitz constant less than 1 (a contraction), then there exists a unique $x_0 \in M$ s.t. $f(x_0) = x_0$.*

Proof. Since X is complete and M is closed, it suffices to prove that the sequence given by $(a_n) := f^n(x)$ is Cauchy. The limit must lie in M as otherwise M is not equal to its closure. By hypothesis there exists $C \in [0, 1)$ s.t.

$$\left|f^2(x) - f(x)\right| < C \cdot |f(x) - x| \qquad \forall\, x \in M$$

This gives the approximation

$$|f^m(x) - x| < \sum_{k=0}^{m-1} C^k\, |f(x) - x| < \left(\sum_{k=0}^{\infty} C^k\right) \cdot |f(x) - x| = \frac{1}{1 - C} \cdot |f(x) - x|$$

which implies that $f^m(x)$ is bounded for all m. Further for $m < n$ we have

$$|f^n(x) - f^m(x)| \le \sum_{k=m}^{n-1} \left|f^{k+1}(x) - f^k(x)\right| < \sum_{k=m}^{n-1} C^k\, |f(x) - x| < C^m \left(\sum_{k=0}^{\infty} C^k\, |f(x) - x|\right) = \frac{C^m}{1 - C}\, |f(x) - x|$$

which can be arbitrarily small via specifying m to be arbitrarily large. Therefore (a_n) admits a limit a. Then

$$a = \lim_{n \to \infty} a_n = \lim_{n \to \infty} f(a_n) = f(\lim_{n \to \infty} a_n) = f(a)$$

since f is continuous (as it is Lipschitz continuous), it commutes with taking the limit.

Uniqueness follows from the fact that if there exists $x_0 \neq x_1$ s.t. $f(x_0) = x_0, f(x_1) = x_1$, then $|x_0 - x_1| = |f(x_0) - f(x_1)|$ which contradicts with the hypothesis. $\square$

Lemma 1.4. *Let X be a normed vector space, with $L \in \{L \in \mathrm{GL}(X) \mid \|L\| < 1\}$ then $(\mathrm{Id} - L) \in \mathrm{GL}(X)$.*

Proof. Consider the series $S = \sum_{i=k}^{\infty} L^k$. This converges as

$$\|S\| \leq \sum_{i=k}^{\infty} \|L\|^k = \frac{1}{1 - \|L\|}$$

which is finite and well-defined as $\|L\| \in [0, 1)$. Then consider

$$\lim_{n \to \infty} (\mathrm{Id} - L) \cdot S = \mathrm{Id} - \lim_{n \to \infty} L^n = \mathrm{Id}$$

as $\|L\| < 1$. $\square$

Proof of Theorem 1.1. Without loss of generality, assume $Df|_{x_0} = \mathrm{Id}$; and $x_0 = 0, f(x_0) = 0$. These are done via applying an affine transformation, which preserves the property of being C^1-invertible.

Consider the approximation of f near $x_0 = 0$: let $g(x) = f(x) - x$. Since $g(0) = 0$ and both f and Id are continuous, there exists r s.t. $\overline{B_r(x)} \subseteq \Omega$, and $\|Dg|_x\| \leq \frac{1}{2}$ for all $x \in \Omega$. This implies that g is a contraction in $\overline{B_r(0)}$, as

$$g(x) - g(0) = \int_0^x Dg|_t \, \mathrm{d}t = \int_0^1 Dg|_{xt} \, x \mathrm{d}t = x \cdot \int_0^1 Dg|_{xt} \, \mathrm{d}t \implies \|g(x)\| \leq \sup_{t \in [0,1]} \|Dg|_{xt}\| \cdot \|x\| = \frac{1}{2} \|x\|$$

This implies that f maps surjectively onto $\overline{B_{r/2}(0)}$: consider $h := y + g(x)$, which is a map from $\overline{B_r(0)}$ to itself. h is also a contraction as

$$\|h(x_1) - h(x_2)\| = \|g(x_1) - g(x_2)\| \leq \frac{1}{2} \|x_1 - x_2\|$$

Further the space on which h acts is closed.[1] This implies that there exists a fixed point x_0 for h; but this implies a preimage:

$$x_0 = y + x_0 - f(x_0) \implies f(x_0) = y$$

Then f is a bijection on $\Omega_1 \to f(\Omega_1)$, where $\Omega_1 = f^{-1} B_{r/2}(0) \cap B_r(0)$. The continuity of f, and the fact that g is a contraction is preserved through restriction of domain. Therefore we have

$$\|x_1 - x_2\| = \|(f(x_1) - f(x_2)) - (g(x_1) - g(x_2))\|$$
$$\leq \|f(x_1) - f(x_2)\| + \|g(x_1) - g(x_2)\|$$
$$\leq \|f(x_1) - f(x_2)\| + \frac{1}{2} \|x_1 - x_2\|$$

which implies that

$$\frac{1}{2} \|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\| \implies \|x_1 - x_2\| = \left\|f^{-1}(y_1) - f^{-1}(y_2)\right\| \leq 2 \|y_1 - y_2\| \tag{$*$}$$

---

[1] The point of taking the closure of $B_r(0)$ in the construction is to apply the contraction mapping principle here, as it requires the space to be closed.

i.e. y is continuous. It is safe to take the inverse as f is a bijection in the domain specified.

Now prove the theorem. By the setup we have $\left\| \mathrm{Id} - Df|_{x_1} \right\| = \left\| Dg|_{x_1} \right\| \leq \frac{1}{2}$ for all $x_1 \in \Omega$, which by Lemma 1.4 implies that $Df|_x$ is invertible for all such x. Since f is differentiable, there exists $Df|_{x_1}$ s.t.

$$f(x) - f(x_1) = Df|_{x_1}(x - x_1) + o(\|x - x_1\|)$$

for $x \in B_\eta(x_1)$ for some η. This implies that

$$(Df|_{x_1})^{-1}(f(x) - f(x_1)) = (x - x_1) + o(\|x - x_1\|)$$

Further restrict $x_1 \in \Omega_1$ where we have the linear bound $(*)$. After substituting the variables we have

$$(Df|_{f^{-1}(y_1)})^{-1}(y - y_1) = (f^{-1}(y) - f^{-1}(y_1)) + o(\|y - y_1\|)$$

which gives by the definition of derivative

$$(Df|_x)^{-1} = Df^{-1}|_{f(x)}$$

Now consider the chain of functions

$$y \xrightarrow{f^{-1}} f^{-1}y \xrightarrow{Df} Df|_{f^{-1}y} \xrightarrow{(-)^{-1}} (Df|_{f^{-1}y})^{-1}$$

where Df is by hypothesis continuous; and f^{-1} is continuous according to the proof above. This implies that the composition, which is the derivative of the inverse function, is continuous, i.e. f is C^1-invertible. $\qquad\square$

This greatly simplifies the proof of the Implicit Function Theorem:

Proof of Theorem 1.2. Consider the map $\Phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$, $(x, y) \mapsto (x, F(x, y))$. Since $\det DF = \det D\Phi \neq 0$ by Laplace expansion, which implies that $D\Phi$ is invertible. This allows application of Inverse Function Theorem, which implies that there exists $\Psi = \Phi^{-1}$ taking the form of $(x, y) \mapsto (x, G(x, y))$. Now consider specifically

$$(x, 0) = \Phi \circ \Psi(x, 0) = \Phi(x, G(x, 0)) = (x, F(x, G(x, 0)))$$

which implies that there exists $g := G(-, 0)$ satisfying the requirement. Since the inverse function is uniquely determined, such g is also uniquely defined. $\qquad\square$

2 Preliminary Topology

Definition 2.1 (Topological Space)**.** *A **topological space** is a pair $(X, \mathcal{O})$, where X is an arbitrary set, and $\mathcal{O}$ a class of subsets of X, which satisfies the follows:*

- *$\emptyset \in \mathcal{O}$, and $X \in \mathcal{O}$.*

- *Let $(\Omega_i)_{i \in I}$ be a (possibly infinite) class of subsets of X, then $\bigcup_{i \in I} \Omega_i \in \mathcal{O}$.*

- *Let $(\omega_i)_{i \in I}$ be a finite class of subsets of X, then $\bigcap_{i \in I} \omega_i \in \mathcal{O}$.*

Elements of $\mathcal{O}$ are called the **open sets**. $\mathcal{O}$ (as a class) gives the topology of X.

Remark 2.2. If the topology of X is clear, one often denotes the space by simply X.

Definition 2.3 (Hausdorff Space). *A topological space $(X, \mathcal{O})$ is a **Hausdorff space** if for all $x, y \in X$ there exists $\Omega_x, \Omega_y \in \mathcal{O}$ s.t. $x \in \Omega_x, y \in \Omega_y$; and $\Omega_x \cap \Omega_y = \emptyset$.*

Example 2.4. Choice of topology is very important. Consider the following examples:

- Consider $(\mathbb{R}, \mathcal{O})$ where $\Omega \in \mathcal{O}$ if and only if for all $x \in \Omega$, there exists $\varepsilon \in \mathbb{R}$ s.t. $(x - \varepsilon, x + \varepsilon) \subseteq \Omega$. This space is Hausdorff as one could choose $\varepsilon_x = \varepsilon_y = |x - y|/4$. This also illustrates why the intersection must be finite, as otherwise one could construct a converging sequence of ε, whose corresponding class of subsets intersecting to a closed interval.

- Consider $(\mathbb{R}, \mathcal{O})$ to be the trivial topology, where $\mathcal{O} := \{\mathbb{R}, \emptyset\}$. Then the only subset in $\mathcal{O}$ containing x and y is $\mathbb{R}$, which implies that the space is not Hausdorff.

Definition 2.5 (Basis of a Topology). *Let $(X, \mathcal{O})$ be a topology space. Then a subset $\mathcal{B} \subseteq \mathcal{O}$ is a **basis of topology** $\mathcal{O}$ if for any $\Omega \in \mathcal{O}$ it can be expressed as union of elements in $\mathcal{B}$. If there exists such $\mathcal{B}$ s.t. it is countably infinite, then X admits a **countable basis of topology**.*

Proposition 2.6. *$(\mathbb{R}, \mathcal{O})$ with topology defined as in the first case in the example above admits a countable basis of topology.*

Proof. First consider $\mathcal{B} := \{(a, b) \mid a < b, a, b \in \mathbb{R}\}$. This gives a basis for the topology of $(\mathbb{R}, \mathcal{O})$ in the sense of the first case above:

- *Any union of elements in $\mathcal{B}$ is open.* Denote such union to be Ω. By construction, for $x \in \mathbb{R}$ s.t. $x \in \Omega$, there exists $a_x < b_x$ s.t. $x \in (a_x, b_x)$. Then there exists $\varepsilon = \min\{x - a_x, b_x - x\}/2$ that satisfies the definition.

- *Any open subset of $\mathbb{R}$ can be expressed as a union of open intervals.* Recall that the union can be infinite. Now consider for Ω an open interval

$$\Omega = \bigcup_{x \in \Omega} (x - \varepsilon_x, x + \varepsilon_x)$$

  where for each x, ε_x is the corresponding radius s.t. the definition is satisfied; and for all $x \in \Omega$, $(\mathbb{R} \setminus \Omega) \cap (x - 2\varepsilon_x, x + 2\varepsilon_x) \neq \emptyset$. By definition for all $x \in \Omega$ there exists such an interval that x is in it.

Then consider $\mathcal{B}' := \{(a, b) \mid a < b, a, b \in \mathbb{Q}\}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for any $x' \in \mathbb{R} \cap \Omega$ there exists a sequence (x_n) s.t. $\lim_{n \to \infty} x_n = x'$. Choose the sequence s.t. $|x_n - x'| > 2|x_{n+1} - x'|$. Suppose that for all n, $x' \notin (x_n - \varepsilon_{x_n}, x_n + \varepsilon_{x_n})$. Then there exists some n_0 s.t. $(x_{n+1} - \varepsilon_{x_{n+1}}, x_{n+1} + \varepsilon_{x_{n+1}}) \subsetneq (x_n, x')$ assuming $x_n < x'$ without loss of generality, which is a contradiction.

Since $\mathbb{Q} \simeq \mathbb{N} \times \mathbb{N}$ which is countable, $\mathbb{Q} \times \mathbb{Q}$ is also countable, indicating that β' gives a countable basis of topology on $\mathbb{R}$. $\qquad\square$

Remark 2.7. For $\mathbb{R}^n$, the standard topology is defined as where a set Ω is open if and only if for every $x \in \Omega$ there exists $\varepsilon > 0$ s.t. $B_\varepsilon^n(x) \subseteq \Omega$.

Corollary 2.8. By the same proof, and considering open intervals as 1-balls on $\mathbb{R}^1$, this gives that $\mathbb{R}^n$ with the standard topology admits a countable basis of topology.

Definition 2.9 (Induced Topology). *Let $(X, \mathcal{O}_X)$ be a topological space, and $Y \subseteq X$. Then there exists a definition for $\mathcal{O}_Y$ where $\Omega' \in \mathcal{O}_Y$ if and only if there exists $\Omega \in \mathcal{O}_X$ s.t. $\Omega' = \Omega \cap Y$. This is the **induced topology** on Y.*

Definition 2.10. *Let $(X_1, \mathcal{O}_1)$, $(X_2, \mathcal{O}_2)$ be topological spaces. Then*

- *A map $f : X_1 \to X_2$ is **continuous** if for all $\Omega \in \mathcal{O}_2$, $f^{-1}(\Omega) \in \mathcal{O}_1$.*

- *A map $f : X_1 \to X_2$ is **homeomorphic** if it is invertible, and both f and f^{-1} are continuous.*

Example 2.11. The map $f : [0, 2\pi) \to S^1, x \mapsto (\cos x, \sin x)$ is not homeomorphic, as for the arcs wrapped around the origin (e.g. $f([0, \pi/6) \cup (11\pi/6, 2\pi))$) the image is open, but the interval itself is not open.

Definition 2.12 (Diffeomorphism)**.** *Let $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ be open sets. Then a homeomorphism $f : \Omega_1 \to \Omega_2$ is a **diffeomorphism** if both f and f^{-1} are differentiable. Similarly one can consider C^k-diffeomorphism for f and f^{-1} being k-times differentiable.*

Proposition 2.13. *If f is a diffeomorphism which is n-times differentiable, then f^{-1} is also n-times differentiable.*

Proof. Proceed to prove this via induction on k:

- *Base case.* For $k = 1$, this is by the definition of diffeomorphisms.

- *Inductive step.* Suppose that this is proven for $k = m$. For $k = m + 1 \leq n$, since f is C^n-differentiable and locally injective everywhere, by inverse function theorem this gives

$$g'(x) = (f'(g(x)))^{-1}$$

where by hypothesis f' and g are both C^m-differentiable, and therefore so is their composition and g'.

This gives g' is C^{n-1}-differentiable, which gives g being C^n-differentiable. $\qquad\square$

Remark 2.14. A differentiable homeomorphism is not necessarily a diffeomorphism, as f is not necessarily locally injective everywhere. Consider $f : x \mapsto x^3$ whose inverse is not differentiable at 0. f is not locally injective at 0.

3   Differentiable Manifolds

Definition 3.1 (Manifold)**.** *An n-**dimensional topological manifold** (n-**manifold**) M is a Hausdorff space with a countable basis of topology s.t. for every $p \in M$ there exists an open $U(p)$ s.t. there exists a homeomorphism $\varphi_p : U(p) \to \Omega \subseteq R^n$.*

Definition 3.2 (Chart; Atlas; Compatibility)**.** *For the homeomorphisms defined as above, the pair $(U(p), \varphi_p)$ is an n-**dimensional (coordinate) chart**. An **atlas** is a set of charts whose domains $U(p)$ covers the whole space M, i.e. for all $r \in M$, there exists $p \in M$ s.t. the corresponding chart $(U(p), \varphi_p)$ satisfies $r \in U(p)$.*

*Given two n-charts (U, φ) and (V, ψ), a **change of chart** between the charts is a map*

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V) \in \{f \mid f : \mathbb{R}^n \to \mathbb{R}^n\}$$

*These two charts are **compatible** if the change of chart is a C^∞-diffeomorphism. An atlas where all charts are compatible is a C^∞-atlas.*

Definition 3.3 (Smooth Manifold)**.** *An n-manifold is **smooth (differentiable)** if it admits a maximal C^∞-atlas (as a set).*

Remark 3.4. Since the atlas is required to be C^∞, it suffices to consider the equivalence class via specifying an atlas whose domains cover the whole manifold, and consider the atlas thus spanned. There exists manifolds that are not smooth (i.e. does not admit a C^∞-atlas) if for all valid charts the maximal set of charts spanned by it does not cover the whole manifold.

Under the context of smooth manifolds, one will not consider atlases that are not C^∞, but instead consider this manifold as not having an atlas.

The followings studies a particular example of manifolds:

Definition 3.5 (Projective Space). *Consider the equivalence relation for $x, y \in \mathbb{R}^{n+1}$ where $x \sim y$ if and only if there exists $\lambda \in \mathbb{R}$ s.t. $\lambda x = y$. The n-**dimensional real projective space** is defined as $\mathbb{R}^{n+1}/(\sim) =: P(\mathbb{R}^{n+1})$, which is often denoted as $\mathbb{RP}^n$.*

Proposition 3.6. $\mathbb{RP}^n$ *is a smooth n-manifold.*

Proof. To specify the manifold structure it suffices to give the topology together with an atlas.

First consider the <u>quotient topology</u> of the projective space: let $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{RP}^n$ be the quotient map specified by the equivalence class each point represents. Define $\Omega' \subseteq \mathbb{RP}^n$ to be open if and only if $f^{-1}(\Omega')$ is open. The topology is indeed well-defined as π commutes with both intersection and union of sets. This also implies that $\mathbb{RP}^n$ admits a countable basis of topology, as by Corollary 2.8 $\mathbb{R}^n$ is compact for all finite n.

Further consider the charts given by

$$\varphi_i : \mathbb{RP}^n \to \mathbb{R}^n, \qquad \overline{(x_1, \cdots, x_{n+1})} \mapsto \frac{1}{x_i}(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$$

where φ_i is defined for all $x \in X_i := \{x \in \mathbb{RP}^n \mid \forall \overline{(x_1, \cdots, x_{n+1})} \in [x], x_i \neq 0\}$. This is clearly open as for all $\tilde{x} \in X_i$, $B_{\tilde{x}_i/2}(\tilde{x}) \subseteq X_i$. Domains of φ_is cover the whole space, as for any $y \in \mathbb{RP}^n$, every representative $\tilde{y}$ of y must have at least some field $\tilde{y}_j$ that is nonzero. Then it is covered by the domain of φ_j. Change of charts is linear according to the construction above, and therefore is smooth. $\square$

Remark 3.7. $\mathbb{RP}^n$ is further compact, as it can be considered as $S^n/(\sim)$ where points sharing the same diameter are considered to be equivalent. The quotient topology implies that $\pi : \mathbb{RP}^n \to S^n/(\sim)$ is continuous. Therefore, $\mathbb{RP}^n$ is compact as S^n is.

Notice that continuous maps preserve compactness. Let $f : M \to N$ be continuous. Then every open cover of N corresponds to an open cover of M, which by compactness of M gives a finite sub-cover U. Then $f(U)$ gives a finite open cover of N.

Definition 3.8 (Submanifold). *Let N be a smooth n-manifold, and $\mathcal{A}$ be its atlas (which is C^∞). Then an m-**dimensional submanifold** M of N is a subset M of N s.t. for all p there exists a chart (U, φ_p) which is in some maximal atlas of N s.t.*

$$M \cap U(p) = \varphi^{-1}(\varphi(U) \cap (\mathbb{R}^m \times \{0\}))$$

*Then $(U \cap M, \varphi_p|_M)$ is called the **adapted chart**; and $(n - m)$ is the **codimension** of M in N.*

Remark 3.9. Every submanifold is in itself a manifold, as restriction of topology preserves Hausdorff property and a countable basis of topology.

]

Definition 3.10 (Smooth Maps)**.** *Consider smooth maps in general between smooth manifolds:*

- *Let $(M, \mathcal{A})$ and $(N, \mathcal{A}')$ be two smooth manifolds. Then $F : M \to N$ is **smooth** if for $(U, \varphi_p) \in \mathcal{A}$ and $(V, \varphi'_{F(p)}) \in \mathcal{A}'$ the map*

$$\varphi'_{F(p)} \circ F \circ \varphi_P^{-1} : \varphi_P(U \cap F^{-1}(V)) \to \varphi'_{F(p)}(F(U) \cap V)$$

  is smooth as an endomorphism on $\mathbb{R}^n$.

- *Let $I \subseteq \mathbb{R}$ be an interval. Then a **smooth curve** on M is a curve with a parameterization $\gamma : I \to M$ that is smooth (considered as a map between manifolds).*

- *$f : M \to \mathbb{R}$ is a **smooth function** if f is smooth as a map between manifolds. Smooth functions on M are denoted by $C^\infty(M)$.*

Definition 3.11. *If $F : M \to N$ is smooth, with F^{-1} exists and is smooth, then F is a **diffeomorphism**, with M and N **diffeomorphic**. Diffeomorphisms that are endomorphisms on M are denoted as $\mathrm{Diff}(M)$, or $\mathrm{Diffeo}(M)$.*

In the next section, more details of smooth manifolds will be discussed, which will lead to natural (and general) construction of submanifolds.

4   Tangent Space and the Differential

It is then immediate to try to approximate functions between manifolds, just as those over real vector spaces. This requires some constructions beforehand:

Definition 4.1 (Parameterized Surface)**.** *A **(smooth) parameterized m-surface** in $\mathbb{R}^n$ is a subset $\mathcal{S} \subseteq \mathbb{R}^n$ s.t. there exists a locally bijective, continuously differentiable map $\sigma : \Omega \to \mathcal{S}$ for $\Omega \subseteq \mathbb{R}^m$ s.t. $\mathrm{rank}\, D\sigma|_x = m$ almost everywhere. For x satisfying such conditions, $f(x) \in \mathcal{S}$ is called a **regular value**.*

Definition 4.2 (Tangent Vector, Tangent Space (in $\mathbb{R}^n$))**.** *Let $\mathcal{S} \subseteq \mathbb{R}^n$ be a smooth parameterized m-surface, together with the parameterization $\sigma : \mathbb{R}^m \supseteq \Omega \to \mathcal{S}$. The k-**th tangent vector at** $p \in \mathcal{S}$ is given by*

$$t_k(p) = \frac{\partial}{\partial x_k} \begin{pmatrix} \sigma_1(x) \\ \vdots \\ \sigma_n(x) \end{pmatrix} \Bigg|_{x = \sigma^{-1}(p)}$$

*and the **tangent space** to $\mathcal{S}$ at p is defined as*

$$T_p\mathcal{S} := \mathrm{span}\{t_1(p), \cdots, t_m(p)\}$$

*with the elements of the tangent space **tangent vectors** to $\mathcal{S}$ at p.*

It is then natural to extend the definition to submanifolds:

Definition 4.3 (Tangent Space (as Submanifolds))**.** *Let $M \subseteq \mathbb{R}^n$ be a m-manifold. Let $p \in M$ together with a chart (U, φ) satisfying that $p \in U$. Then the **tangent space of M at** p is defined as*

$$T_pM := (\,D\varphi|_p)^{-1}(\mathbb{R}^m \times \{0\})$$

Remark 4.4. Notice that the tangent space defined as above is independent of the choice of chart. Inverse Function Theorem gives $\left(D\varphi|_p \right)^{-1} = D\varphi^{-1}\big|_{\varphi(p)}$. Let (V, ψ) be another chart s.t. it is compatible with (U, φ). This allows expanding the chart using the chain rule:

$$\operatorname{im} D(\varphi^{-1})\big|_{\varphi(p)} = \operatorname{im} D(\psi^{-1} \circ \psi \circ \varphi^{-1})\big|_{\varphi(p)}$$
$$= \operatorname{im} D\psi^{-1}\big|_{\psi(p)} \cdot D(\psi \circ \varphi^{-1})\big|_{\varphi(p)}$$

Since both ψ and φ are invertible, $\psi \circ \varphi^{-1}$ are locally bijective, which implies that $\operatorname{im} D(\varphi^{-1})\big|_{\varphi(p)} = \operatorname{im} D(\psi^{-1})\big|_{\psi(p)}$

Remark 4.5. Notice that the definition above requires the existence of an "ambient space" around the manifold ($\mathbb{R}^n$), as this characterizes elements in $T_p M$ (as elements in $\mathbb{R}^n$).

Definition 4.6 (Coordinate Function, Directional Derivative). *The **coordinate function** of standard coordinates x^i is the extraction of i-th coordinate:*

$$x^i : \mathbb{R}^n \to \mathbb{R}, \qquad (x_1, \dots, x_n) \mapsto x_i$$

*This immediately gives the **coordinate function of chart** (U, φ) to be*

$$\varphi^i : M \to R, \qquad \varphi^i := x^i \circ \varphi$$

*Then, for a function $f : M \to \mathbb{R}^n$, one can naturally define the **directional derivative** (in the only way that is syntactically true):*

$$\frac{\partial}{\partial \varphi^i}\bigg|_p (f) = \frac{\partial f}{\partial \varphi^i}\bigg|_p := \frac{\partial (f \circ \varphi^{-1})}{\partial x^i}\bigg|_{\varphi(p)}$$

where the right-most term is simply the partial differentiation of functions between Euclidean spaces.

The following forms an "standalone" definition, which turns out to be compatible with the definition above under the context of submanifolds.

Definition 4.7 (Tangent Vector (Standalone)). *Let M be a smooth n-manifold, with $p \in M$. Then*

*i) An **algebraic tangent vector (as derivatives)** is defined as a linear map $\partial : C^\infty(M) \to \mathbb{R}$ where $C^\infty(M)$ is defined as the smooth functions $M \to \mathbb{R}$, which satisfies the product rule (Leibniz rule):*

$$\forall f, g \in C^\infty(M), \qquad \partial(f \cdot g) = (\partial f) \cdot g + f \cdot (\partial g)$$

*ii) A **physical tangent vector** associated to a chart (U, φ) where $p \in \varphi$, which is essentially a transformation into $\mathbb{R}^n$ (as in the previous case where the tangent space is treated as a submanifold of $\mathbb{R}^n$), is defined as a vector $v_\varphi \in \mathbb{R}^n$, satisfying the compatibility of change of charts: for all charts (V, ψ) s.t. $p \in V$,*

$$v_\psi = D(\psi \circ \varphi^{-1})\big|_{\varphi(p)} \cdot v_\varphi$$

Expressed in each coordinate, the transformation essentially becomes the directional derivative

$$v_\psi^i = \sum_{k=1}^n \frac{\partial(\psi \circ \varphi^{-1})}{\partial x^k}\bigg|_{\varphi(p)} (v_\varphi^k) = \sum_{k=1}^n \frac{\partial \psi}{\partial \varphi^k}\bigg|_p (v_\varphi^k)$$

iii) *A **geometric tangent vector**, or a **tangent curve** at p is an equivalence class of smooth curves $\gamma : I \to M$, s.t. $0 \in I \subseteq R$, and $\gamma(0) = p$; and for p supported by the chart (U, φ) the equivalence relation given by*

$$\gamma_1 \sim \gamma_2 \quad \Longleftrightarrow \quad (\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0)$$

It will be shown later that all these three definitions are equivalent.

Remark 4.8. The definition of algebraic tangent vectors is completely unrelated to the charts. The equivalence relation defined in geometric tangent vector is also independent of the chart, as for a different chart (V, ψ) s.t. $p \in V$,

$$
\begin{aligned}
(\psi \circ \gamma_1)'(0) &= (\psi \circ \varphi^{-1} \circ \varphi \circ \gamma_1)'(0) \\
&= (\psi)'(\gamma_1(0)) \cdot (\varphi^{-1})'(\varphi \circ \gamma_1(0)) \cdot (\varphi \circ \gamma_1)'(0) \\
&= (\psi)'(p) \cdot (\varphi^{-1})'(\varphi(p)) \cdot (\varphi \circ \gamma_1)'(0) \\
&= (\psi)'(\gamma_2(0)) \cdot (\varphi^{-1})'(\varphi \circ \gamma_2(0)) \cdot (\varphi \circ \gamma_2)'(0) \\
&= (\psi \circ \gamma_2)'(0)
\end{aligned}
$$

Theorem 4.9. *The definition of previous three tangent vectors are equivalent. That is, given any one of them, the other two can be uniquely determined.*

Proof. It suffices to give the construction of the corresponding tangent vectors:

- i) $\implies$ ii). Given an algebraic tangent vector $\partial : C^\infty(M) \to \mathbb{R}$, for $p \in U \cap V$ where (U, φ) and (V, ψ) are two charts of M, define

$$v_\varphi^i := \partial(\varphi^i)$$

  Check that this satisfies the compatibility requirement:

$$
\begin{aligned}
v_\psi^i = \partial(\psi^i) &= \partial((\psi^i \circ \varphi^{-1}) \circ \varphi) \\
&\overset{1}{=} \sum_{j=1}^n \partial(\psi^i \circ \varphi^{-1} \circ \varphi^j) \\
&\overset{2}{=} \sum_{j=1}^n \left. \frac{\partial(\psi^i \circ \varphi^{-1})}{\partial x^i} \right|_{\varphi(p)} v_\varphi^j
\end{aligned}
$$

  where equality 1 holds by the linearity of ∂; and equality 2 is essentially the chain rule applied to derivation, which is determined by the requirement that the Leibniz rule holds. This is exactly the form required.

- ii) $\implies$ i). For any instance v_φ as a physical tangent vector under chart (U, φ), define

$$\partial := \sum_{i=1}^n v_\varphi^i \left. \frac{\partial}{\partial \varphi^i} \right|_p$$

  Check the followings:

  - ∂ *is well-defined, i.e. does not depend on the choice of chart.* Expand using the definition of directional derivatives and

transition between charts of physical tangent vectors, we have for all $f \in C^\infty(M)$,

$$\partial(f) := \sum_{i=1}^{n} v_\varphi^i \cdot \left.\frac{\partial}{\partial \varphi^i}\right|_p (f) = \sum_{i=1}^{n} \left.\frac{\partial(f \circ \varphi^{-1})}{\partial x^i}\right|_{\varphi(p)} (f)$$

$$= \sum_{i=1}^{n} \left.\frac{\partial(f \circ (\varphi^i)^{-1})}{\partial x^i}\right|_{\varphi(p)} (f)$$

$$\overset{1}{=} \sum_{i=1}^{n} \left(\left(\sum_{j=1}^{n} \left.\frac{\partial(\varphi^i \circ \psi^{-1})}{\partial x^j}\right|_{\psi(p)} v_\psi^j \right) \cdot \left.\frac{\partial(f \circ (\varphi^i)^{-1})}{\partial x^i}\right|_{\varphi(p)} (f) \right)$$

$$\overset{2}{=} \sum_{j=1}^{n} \left(v_\varphi^j \left(\sum_{i=1}^{n} \left.\frac{\partial(\varphi^i \cdot \psi^{-1})}{\partial x^j}\right|_{\psi(p)} \right) \cdot \left.\frac{\partial(f \circ (\varphi^i)^{-1})}{\partial x^i}\right|_{\varphi(p)} (f) \right)$$

$$\overset{3}{=} \sum_{j=1}^{n} \left(v_\varphi^j \cdot \left.\frac{\partial(\varphi^j \circ \psi^{-1})}{\partial x^j}\right|_{\psi(p)} \cdot \left.\frac{\partial(f \circ (\varphi^i)^{-1})}{\partial x^j}\right|_{\varphi(p)} (f) \right)$$

$$= \sum_{j=1}^{n} v_\varphi^j \cdot \left.\frac{\partial(f \circ \psi^{-1})}{\partial x^j}\right|_{\psi(p)} (f) = \sum_{i=1}^{n} v_\psi^i \cdot \left.\frac{\partial}{\partial \psi^i}\right|_p (f)$$

where

1. Equality 1 holds by the definition of the transition between physical tangent vectors.

2. Equality 2 holds by changing the summation between i and j. Notice that $\left.\frac{\partial(\varphi^i \cdot \psi^{-1})}{\partial x^j}\right|_{\psi(p)}$ is a scalar, while applying it to v_ψ^j is simply multiplication.

3. Equality 3 holds as the variables x^i and x^j are independent for $i \neq j$. Therefore the summation can be simplified to be only with one variable.

   – *Leibniz rule holds.* This is simply the product rule of normal derivatives.

- *ii)* $\implies$ *iii).*

$\square$