Differential Geometry

Chin-Lung Wang

ABSTRACT. This is the preliminary version of my course notes in the fall term of 2006 at NCU and 2012 at NTU. The aim is to provide basic concepts in differential geometry for first year graduate students as well as advanced undergraduate students.

Contents

Chapter 1

Differentiable Manifolds

We start by defining (finite dimensional) differentiable manifolds M and their tangent bundles TM, and proving various basic results to establish the categorical concept of manifolds. Notably for a differentiable map $f : M \to N$ we characterizes immersions and submersions through its tangent map $df_p : T_pM \to T_{f(p)}N$.

Manifolds are locally Euclidean spaces such that the local coordinates transforms smoothly. Thus most of the technical results are based on their counterparts in coordinate charts (Calculus) and a good way to glue them together (partitions of unity). We give proofs of Whitney' imbedding theorem and Sard's theorem on critical values to illustrate such a principle.

A more rigid type of argument to establish global results on manifolds are through "analytic continuations" based on differential equations. In this beginning chapter we will only discuss ordinary differential equations with emphasized on the uniqueness and smooth dependence of the solutions on the initial conditions. This leads to the notion of integral curves and flows which allows us to define Lie derivatives of vector fields, which is fundamental throughout this course.

1. The category of C^k manifolds

Definition 1.1. A *topological manifold* M is a topological space which is (1) locally Euclidean (2) Hausdorff and (3) second countable.

Here are some explanations of these concepts:

(1) M is *locally Euclidean* if for each point $p \in M$ there is a open neighborhood $U \ni p$ which is homeomorphic to an open set in $\mathbb{R}^d$ for some $d \in \mathbb{N}$. Let

$$\varphi : U \to \varphi(U) \subset \mathbb{R}^d$$

be such a homeomorphism. The components $x^i : U \to \mathbb{R}$ of φ are called the coordinate functions and the pair (U, φ) is called a (coordinate) *chart* of M at p. It is customary to identify ϕ with the (column vector) coordinate function

$$\mathbf{x} = (x^1, \ldots, x^d)^t.$$

(2) M is *Hausdorff* if for any $p \neq q$ in M there are neighborhood $U \ni p$, $V \ni q$ such that $U \cap V = \emptyset$.

(3) M is *second countable* if there is countable basis for its topology. Recall that a basis is a collection of open subset such that any open set can be written as a (possibly infinite) union of certain constituents from that collection.

Exercise 1.1. Show that $\mathbb{R}^d$ (with the standard Euclidean topology) is a manifold by finding an explicit countable basis.

It is not a priori clear why condition (3) should be there. A possible reason goes as follows: if a topological space M is Hausdorff and second countable, then any subset $S \subset M$ with the induced topology is also Hausdorff and second countable. In particular, any locally Euclidean subset in $\mathbb{R}^d$ is a manifold. Conversely, we will prove later that any manifold as defined above is indeed a subspace in $\mathbb{R}^d$ (the Whitney Imbedding Theorem, theorem 1.22). Hence, we see that above abstract definition of manifolds does not really lead to anything outside Euclidean spaces.

Given a manifold M and two charts (U_a, ϕ_a) in $\mathbb{R}^{d_a}$ and (U_b, ϕ_b) in $\mathbb{R}^{d_b}$ with $U_a \cap U_b \neq \emptyset$, we form the coordinate transition function

$$\phi_{ab} := \phi_a \circ \phi_b^{-1} : \phi_b(U_a \cap U_b) \to \phi_a(U_a \cap U_b)$$

which is a homeomorphism. It is intuitively clear that we should have $d_a = d_b$, which will be the dimension of M. However, the only known proofs are by no means elementary, except in one case:

Exercise 1.2. Let $\mathbb{R}^{d_1} \cong \mathbb{R}^{d_2}$ (homeomorphic). If $d_1 = 1$ show that $d_2 = 1$. Investigate the case $d_1 = 2$ and reduce the problem to the Jordan Curve Theorem.

The general case will be outlined later (c.f. Exercise 1.19) by means of certain *approximation theorems* and ideas in *homotopy theory*.

In this course we are mainly interested in *differentiable manifolds* instead of general topological manifolds.

Definition 1.2. We call a collection of charts $\{(U_a, \phi_a)\}_{a\in A}$ a C^k *atlas* of M if (1) the transition functions ϕ_{ab}'s are all C^k mappings for some fixed $k \in \mathbb{N} \cup \{\infty\}$ and (2) $\bigcup_{a\in A} U_a = M$.

Exercise 1.3. For a manifold with a C^k atlas, $k \geq 1$, show that the dimension $d = \dim M$ is well defined on each connected component of M.

When a manifold M is equi-dimensional of dimension d, we usually denote it by M^d, if no confusion with the cartesian product $M \times \cdots \times M$ is likely to occur.

Given a C^k atlas $\{(U_a, \phi_a)\}_{a\in A}$ on M, a chart (U, ϕ) is C^k *related* to it if both the transition functions $\phi \circ \phi_a^{-1}$ and $\phi_a^{-1} \circ \phi$ are C^k for all $a \in A$. It is convenient to add all C^k related charts into a given atlas.

Exercise 1.4. Show that the enlarged collection of charts $\{U_\alpha, \phi_\alpha)\}_{\alpha\in\mathcal{A}}$ also forms a C^k atlas. Moreover, it is a maximal atlas in the sense that any chart which is C^k related to it is already contained in it.

Definition 1.3. A C^k *(differentiable) structure* on M is a maximal atlas of C^k charts. A C^k differentiable manifold is a manifold together with a C^k structure. When a C^k manifold is given, the term *charts of it* will always mean C^k charts.

Formally, the case $k = 0$ is simply a topological manifold. From the definition, there is an immediately natural question whether it is possible to select from all charts a sub-collection which defines a C^1 structure or even a C^k structure for higher k. These are important and highly non-trivial problems in manifold theory. In fact, there are C^0 manifolds which admit no C^1 structures. In later chapters we will address on some of these questions. For the moment, we will only remark that:

(1) A famous theorem of Whitney says that any C^1 manifold indeed admits (contains) C^∞ structures;
(2) the C^2 condition is the minimum requirement to define the notion of *curvature*, a concept introduced by Gauss and Riemann which lead to the birth of modern differential geometry, and will be vastly studied in this course.

Thus in this course, differentiable manifolds will always mean C^∞ manifolds.

A function $f : M^d \to \mathbb{R}$ is C^k *(differentiable)* at $p \in M$ if $f \circ \mathbf{x}^{-1}$ is C^k at $\mathbf{x}(p) \in \mathbb{R}^d$ for one chart $(U, \mathbf{x})$ which contains p. Since

$$f \circ \mathbf{x}_\beta^{-1} = f \circ \mathbf{x}_\alpha^{-1} \circ (\mathbf{x}_\alpha \circ \mathbf{x}_\beta^{-1}),$$

by the definition of C^k structure the notion of C^k is independent of the choices of charts. We denote $C^k(U)$ by the space of functions that are C^k at all points in U.

Likewise a function $f : M^m \to N^n$ between two C^k manifolds is called a C^k *(differentiable) map* if

$$\mathbf{y} \circ f \circ \mathbf{x}^{-1} : \mathbf{x}(f^{-1}(V) \cap U) \subset \mathbb{R}^m \to \mathbf{y}(V) \subset \mathbb{R}^n$$

is C^k for any choice of charts $(U, \mathbf{x})$ on M and $(V, \mathbf{y})$ on N. It is enough to check it for any two special atlas. Denote by $C^k(M, N)$ the space of all such C^k functions. A mapping $f : M \to N$ between two C^k manifolds is a *diffeomorphism* if f^{-1} is well defined and both f and f^{-1} are C^k. This is the notion of isomorphisms in the category of C^k manifolds.

Exercise 1.5. For any C^k manifold M^d and $p \in M$, show that there are charts with $\mathbf{x}(U) = B_0(r)$, the open ball of radius r in $\mathbb{R}^d$, as well as charts with $\mathbf{x}(U) = \mathbb{R}^d$.

Exercise 1.6. Consider $M = \mathbb{R}$ with one chart given by $(\mathbb{R}, \phi)$ where $\phi(t) = t^3$. Show that this defines a C^∞ structure on M. Is M diffeomorphic to $\mathbb{R}$ with the standard C^∞ structure $(\mathbb{R}, \mathrm{id})$?

There could be many C^k structures on a manifold, but it is hard to find non-diffeomorphic ones. The set of equivalence classes of differentiable structures up to diffeomorphism is a delicate object for study, which again will be briefly discussed in later chapters.

2. Cut offs and the partition of unity

Are there any C^∞ functions on a C^∞ manifold M besides the constants? For each charts $(U, \mathbf{x})$ the coordinate functions x^i's are by definition C^∞ on U, but it may not be possible to extend x^i to a C^∞ function on M.

One of the basic principles in differential geometry is try to (1) compute things locally via differential calculus and (2) find a way to patch local information together to get global results. This section establishes the existence of *partitions of unity* which is the simplest tool in this regard.

Recall that a topological space M is paracompact if every open cover $\{U_\alpha\}_{\alpha\in A}$ of it has a locally finite open refinement $\{V_\beta\}_{\beta\in B}$, in the sense that

(1) *Local finiteness*: for each $p \in M$, there is a neighborhood $U \ni p$ such that $V_\beta \cap U = \emptyset$ except possibly for a finite number of V_β's.

(2) *Refinement*: there is a map $\rho : B \to A$ such that $V_\beta \subset U_{\rho(\beta)}$ for all $\beta \in B$. The map ρ may not be injective nor surjective.

A manifold is more than paracompact. In fact we have an easy but important

Lemma 1.4. *Let M be a locally compact topological space which is Hausdorff and second countable (e.g. a manifold), then M is σ compact. Namely, there is a countable sequence of increasing open sets $\{G_i\}_{i\in\mathbb{N}}$ with $\bar{G}_i$ compact, $\bar{G}_i \subset G_{i+1}$ and $M = \bigcup_{i=1}^\infty G_i$.*

PROOF. Let $\{W_i\}_{i\in\mathbb{N}}$ be any given countable basis.

Exercise 1.7. Show that by removing those W_i with noncompact closure $\bar{W}_i$ we still get a basis. (Notice that the Hausdorff condition is needed.)

Thus we may assume that $\bar{W}_i$ is compact for all i.

Let $G_1 = W_1$. The set G_i is constructed inductively: suppose that G_i is constructed. Since $\bar{G}_i$ is compact and covered by W_j's, there is a smallest $j(i) \in \mathbb{N}$ so that

$$\bar{G}_i \subset W_1 \cup \cdots \cup W_{j(i)}.$$

We then define $G_{i+1} = W_1 \cup \cdots \cup W_{j(i)}$. It remains to show that $\bar{G}_{i+1}$ is compact. This follows from

$$\bar{G}_{i+1} \subset \bar{W}_1 \cup \cdots \cup \bar{W}_{j(i)}$$

since a closed set in a (finite union of) compact set is compact. □

Lemma 1.5. *Let $M = \bigcup_{i=1}^{\infty} G_i$ be σ compact. Then every open cover $\{U_\alpha\}_{\alpha \in A}$ has a countable locally finite refinement $\{V_j\}_{j \in \mathbb{N}}$ with $\bar{V}_j$ being compact.*

PROOF. For each $i \in \mathbb{N}$, consider the open annulus

$$S_i := G_{i+1} \setminus \bar{G}_{i-2}.$$

(We put $G_i = \emptyset$ for $i \leq 0$.) Then $\bar{G}_i \setminus G_{i-1}$ is compact and contained in S_i. It is covered by $\{U_\alpha \cap S_i\}_{\alpha \in A}$ hence is covered by a finite number of them. By putting together all these finite open sets we get a countable sequence $\{V_j\}_{j \in \mathbb{N}}$. Each V_j is of the form $U_\alpha \cap S_i$, so

$$\bar{V}_j = \overline{U_\alpha \cap S_i} \subset \bar{S}_i \subset \bar{G}_{i+1}$$

is closed in a compact set. Hence $\bar{V}_j$ is itself compact.

Finally, $\{V_j\}$ is locally finite since if $p \in S_i$ then only those V_j's constructed from S_{i-1}, S_i and S_{i+1} may possibly intersect S_i nontrivially. □

Now we discuss cut off (or bump) functions . Let

$$f(t) = \begin{cases} e^{-1/t} & \text{for} \quad t > 0, \\ 0 & \text{for} \quad t \leq 0. \end{cases}$$

Exercise 1.8. Show that $f \in C^\infty(\mathbb{R})$ and $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$.

The function

$$g(t) = \frac{f(t)}{f(t) + f(1-t)} = \frac{1}{1 + e^{\frac{1}{t} - \frac{1}{1-t}}}$$

is then C^∞ and non-decreasing with $g(t) = 0$ for $t \leq 0$ and $g(t) = 1$ for $t \geq 1$.

The function $h(t) = g(2+t)g(2-t)$ is a cut off function with $h = 1$ on $[-1, 1]$ and $h = 0$ outside $(-2, 2)$. For a higher dimensional version we consider

$$\psi(x^1, \ldots, x^d) = \prod_{i=1}^{d} h(x^i) \in C^\infty(\mathbb{R}^d).$$

Then $\psi = 1$ on $[-1, 1]^d$ and $\psi = 0$ outside $(-2, 2)^d$. Alternatively we may consider the radially symmetric function

$$\psi(x) = h(|x|) \in C^\infty(\mathbb{R})$$

which has $\psi|_{B_0(1)} = 1$ and $\psi|_{\mathbb{R}^d \setminus B_0(2)} = 0$.

In general, for a continuous function f on a topological space M, its support is defined to be

$$\operatorname{supp} f := \overline{\{p \in M \mid f(p) \neq 0\}}.$$

For a closed set $B \subset M$, a cut off function for B is a non-negative continuous function f such that $\operatorname{supp} f = B$. The functions ψ above are special C^∞ cut off functions of standard cube and closed balls.

Definition 1.6. Given an open cover $\{U_\alpha\}_{\alpha \in A}$ of a C^k manifold M, a partition of unity subordinate to $\{U_\alpha\}$ is a countable collection of C^k functions $\{\psi_j\}_{j \in \mathbb{N}}$ on M such that

(1) $0 \leq \psi_j \leq 1$ for all j.
(2) $\{\operatorname{supp} \psi_j\}$ is a locally finite (closed) refinement of $\{U_\alpha\}$.
(3) $\sum_{j \in \mathbb{N}} \psi_j(p) = 1$ for all $p \in M$.

There will be no convergence issue in (3) since by (2) the sum will be a finite sum over a neighborhood of any point p.

Theorem 1.7 (Existence of Partition of Unity). *Let M be a C^k manifold with $\{U_\alpha\}_{\alpha \in A}$ an open cover. Then there is a C^k partition of unity $\{\psi_i\}_{j \in \mathbb{N}}$ subordinate to $\{U_\alpha\}$ with* $\operatorname{supp} \psi_j$ *being compact.*

Without the compact support requirement we may label the partition of unity by the same set A with $\psi_\alpha \not\equiv 0$ for at most a countable subset of A.

PROOF. Let $M = \bigcup_{i=1}^\infty G_i$ as given by Lemma 1.4. We will modify the proof of Lemma 1.5 to construct $\{\psi_j\}$.

For each $i \in \mathbb{N}$, the region $\bar{G}_i \setminus G_{i-1}$ is compact and contained in $S_i = G_{i+1} \setminus \bar{G}_{i-2}$. For each $p \in \bar{G}_i \setminus G_{i-1}$, let $(W_p, \mathbf{x})$ be a chart at p such that $W_p \subset U_\alpha \cap S_i$ for some $\alpha \in A$ and $\mathbf{x}(W_p) = B_0(3)$. Let $V_p = \mathbf{x}^{-1}(\bar{B}_0(2))$. Define a C^k cut off function for $\bar{V}_p \subset W_p \subset U_\alpha$ by

$$\bar{\psi}_p = \begin{cases} \psi \circ \mathbf{x} & \text{on} \quad W_p, \\ 0 & \text{on} \quad M \setminus W_p. \end{cases}$$

There is a finite subcover of the open cover $\{V_p\}$ of $\bar{G}_i \setminus G_{i-1}$. By putting together all such finite open sets for all $i \in \mathbb{N}$, we get the desired locally finite refinement $\{V_j\}_{j \in \mathbb{N}}$ as in Lemma 1.5.

Let $\bar{\psi}_j$ be the corresponding cut off function for V_j. For each $p \in M$, there is a (finite number of) $\bar{\psi}_j$ with $\bar{\psi}_j(p) \neq 0$, hence we may define

$$\psi_j = \frac{\bar{\psi}_j}{\sum_i \bar{\psi}_i} \in C^\infty(M),$$

which clearly satisfies $\sum_j \psi_j = 1$ with $\operatorname{supp} \psi_j = \operatorname{supp} \bar{\psi}_j = \bar{V}_j$ being compact.

For the last statement, for each $\alpha \in A$, we may simply let

$$\psi_\alpha = \sum_{\bar{V}_j \subset U_\alpha} \psi_j.$$

Here $\psi_\alpha \equiv 0$ if no such j exists. The proof is complete. □

Exercise 1.9. Investigate the theorem for the case when $M = \mathbb{R}$ with the open cover being given by a single set $U = M = \mathbb{R}$.

Exercise 1.10. Let A (resp. U) be a closed (resp. open) set in a C^k manifold M with $\bar{A} \subset U$. Show that there exists $f \in C^k(M)$ such that $f|_A \equiv 1$ and $f|_{M \setminus U} \equiv 0$. Is that possible to make $\operatorname{supp} f = A$?

3. Tangent spaces

It is a priori not obvious how to generalized the concept of tangent vectors to manifolds. In fact, this is a challenge problem for C^0 manifolds. We will give two definitions of it for C^k manifolds with $k \in \mathbb{N} \cup \{\infty\}$.

Let us begin with the case of Euclidean spaces. Let $p \in \mathbb{R}^d$ and $X \in \mathbb{R}^d$ be a vector. For a C^1 function f defined near p, we can define the directional derivative

$$Xf := D_X f(p) = \frac{d}{dt} f(p+tX)\big|_{t=0} = \lim_{t \to 0} \frac{f(p+tX) - f(p)}{t}.$$

It is a derivation (first order differential operator) in the sense that for all C^1 functions f, g defined near p, we have

(1) *Linearity*: $X(af + bg) = aXf + bXg$ for $a, b \in \mathbb{R}$, and
(2) *Leibniz rule*: $X(fg) = (Xf)g(p) + f(p)Xg$.

Conversely, it is interesting to see whether a derivation determines a vector. We will see shortly that this is indeed the case for derivations on C^∞ functions.

For a C^k manifold, denote by C_p^k the space of *germs of C^k functions at p*. It consists of functions which are defined on some (open) neighborhood U of p, and two function germs f, g are identified if $f|_U = g|_U$ for some $U \ni p$.

Definition 1.8. Let M be a C^k manifold and $p \in M$. The *Zariski tangent space D_pM* at p is the vector space consisting of all derivations $X : C_p^k \to \mathbb{R}$.

For any chart $(U, \mathbf{x})$ at p, partial derivatives $\partial/\partial x^i|_p$ are examples of tangent vectors: for $f \in C_p^k$,

$$\frac{\partial}{\partial x^i}\Big|_p f := \frac{\partial(f \circ \mathbf{x}^{-1})}{\partial x^i}(\mathbf{x}(p)).$$

The following lemma is a useful substitute of the Taylor expansion especially for functions that has only limited differentiability.

Lemma 1.9. *Let $f \in C^k(B_0(r))$ in $\mathbb{R}^d$. Then*

$$f(x^1, \ldots, x^d) = f(0) + \sum_{i=1}^{d} x^i g(x^1, \ldots, x^d)$$

with $g_i \in C^{k-1}(B_0(r))$ and $g_i(0) = \partial f / \partial x^i(0)$.

PROOF. By the Fundamental Theorem of Calculus,

$$\begin{aligned} f(x) - f(0) &= \int_0^1 \frac{d}{dt} f(tx)\, dt = \sum_{i=1}^{d} \int_0^1 \frac{\partial f}{\partial x^i}(tx) x^i \, dt \\ &= \sum_{i=1}^{d} x^i \int_0^1 \frac{\partial f}{\partial x^i}(tx)\, dt = \sum_{i=1}^{d} x^i g(x), \end{aligned}$$

where $g_i(x) = \int_0^1 \frac{\partial f}{\partial x^i}(tx)\, dt \in C^{k-1}(B_0(r))$ and $g(0) = \frac{\partial f}{\partial x^i}(0)$ as expected. □

Theorem 1.10. *For a C^∞ manifold with $(U, \mathbf{x})$ a chart at p, the partial derivatives form a basis of $D_p M$. Indeed for any $X \in D_p M$,*

$$X = \sum_{i=1}^{d} X(x^i) \frac{\partial}{\partial x^i}\Big|_p .$$

PROOF. Since $X(1) = X(1 \cdot 1) = X(1) \cdot 1 + 1 \cdot X(1) = 2X(1)$, we have $X(1) = 0$, hence $X(a) = aX(1) = 0$ for any constant a. For simplicity of notations we assume that $\mathbf{x}(p) = 0$. Then for any $f \in C_p^\infty$,

$$\begin{aligned} Xf &= X(f - f(p)) = X(\textstyle\sum_i x^i g_i) \\ &= \textstyle\sum_i X(x^i) g_i(p) + x^i(p) X(g_i) \\ &= \textstyle\sum_i X(x^i) \dfrac{\partial (f \circ \mathbf{x}^{-1})}{\partial x^i}(0) \\ &= \textstyle\sum_i X(x^i) \dfrac{\partial}{\partial x^i}\Big|_p f. \end{aligned}$$

This proves the theorem. □

The proof (and the theorem) fails for C^k manifolds if $k < \infty$ because then $g_i \in C^{k-1}$ only and the second line about the Leibniz rule does not make sense. This can be analyzed in purely algebraic terms:

Proposition 1.11. *For any C^k manifold M, $k \in \{0, 1, 2, \cdots, \infty\}$,*

$$D_p M \cong (m_p/m_p^2)^*.$$

Here $m_p = \{f \in C_p^k \mid f(p) = 0\}$ is the maximal ideal at p and m_p^2 consists of all finite sum of products $\sum f_i g_i$ with $f_i, g_i \in m_p$.

PROOF. Let $X \in D_p M$. Then X defines a linear map $X : m_p \to \mathbb{R}$. To show that X induces a map $m_p/m_p^2 \to \mathbb{R}$ we need $X|_{m_p^2} = 0$, which follows from the Leibniz rule readily: for $f_i, g_i \in m_p$,

$$X(\textstyle\sum_i f_i g_i) = \sum_i X(f_i) g_i(p) + f_i(p) X g_i(p) = 0.$$

Conversely, given $\psi : m_p/m_p^2 \to \mathbb{R}$ we claim that $X_\psi f := \psi(f - f(p))$ defines a derivation X_ψ on C_p^k. Indeed,

$$\begin{aligned} X_\psi(fg) &= \psi(fg - f(p)g(p)) \\ &= \psi((f - f(p))(g - g(p)) + (f - f(p))g(p) + f(p)(g - g(p))) \\ &= (X_\psi f)g(p) + f(p) X_\psi g \end{aligned}$$

where we use the fact that $(f - f(p))(g - g(p)) \in m_p^2$. $\square$

Exercise 1.11. Let M be a C^k manifold. Show that

$$\dim D_p M = \dim m_p/m_p^2 = \begin{cases} \dim M & \text{if} \quad k = \infty, \\ \infty & \text{if} \quad k < \infty. \end{cases}$$

(Hint: for $k = 1$, study functions $f(x) = (x^1)^a$ for $1 < a < 2$.)

Let $(U, \mathbf{x})$ and $(V, \mathbf{y})$ be two charts at p, then for any $f \in C_p^\infty$,

$$\begin{aligned} \frac{\partial}{\partial x^i}\Big|_p f &= \frac{\partial(f \circ \mathbf{x}^{-1})}{\partial x^i}(\mathbf{x}(p)) = \frac{\partial(f \circ \mathbf{y}^{-1} \circ \mathbf{y} \circ \mathbf{x}^{-1})}{\partial x^i}(\mathbf{x}(p)) \\ &= \sum_j \frac{\partial(f \circ \mathbf{y}^{-1})}{\partial y^j}(\mathbf{y}(p)) \frac{\partial y^j}{\partial x^i}(\mathbf{x}(p)) = \sum_j \frac{\partial y^j}{\partial x^i}(\mathbf{x}(p)) \frac{\partial}{\partial y^j}\Big|_p f. \end{aligned}$$

Thus two vector representations $\sum_i a^i \frac{\partial}{\partial x^i}\Big|_p$ and $\sum_j b^i \frac{\partial}{\partial y^j}\Big|_p$ on two charts correspond to *the same vector* precisely when their coefficients satisfy the transformation rule

$$b^j = \sum_i a^i \frac{\partial y^j}{\partial x^i}(\mathbf{x}(p)).$$

This condition only requires C^1 structure, so we make the

Definition 1.12. The *tangent space* T_pM for a C^k manifold M with $k \geq 1$ consists of compatible systems of vectors in each chart at p which satisfy the transformation rule.

4. Tangent maps

From now on, we work in the C^∞ category unless specified otherwise. In particular $T_pM = D_pM$ and tangent vectors are precisely derivations. Given a C^∞ map $f : M \to N$ and $p \in M$, we define the tangent map $df_p : T_pM \to T_{f(p)}N$ via

$$(df_pX)h = X(h \circ f),$$

where $X \in T_pM$ and $h \in C^\infty_{f(p)}$.

The tangent map is indeed the generalization of derivative of a map in Calculus and is also denoted by Df_p, $Df(p)$, $df(p)$, f_{*p}, and perhaps the most commonly used $f'(p)$. It is the linearization (first–order approximation) of the original map.

As in Calculus, df_p is linear and satisfies the *chain rule*. Namely for $g : N \to S$ be another C^∞ map we have

$$d(g \circ f)_p = dg_{f(p)} \circ df_p.$$

The proof in Calculus is a bit tricky, but the proof now is completely formal: for $X \in T_pM$ and $h \in C^\infty_{gf(p)}$,

$$(d(g \circ f)_pX)h = X(h \circ g \circ f) = (df_pX)(h \circ g) = dg_{f(p)}(df_pX)h.$$

Exercise 1.12. Show that in charts $(U, \mathbf{x})$ at $p \in M$ and $(V, \mathbf{y})$ at $f(p) \in N$, df_p is represented by $d\tilde{f}_{\mathbf{x}(p)}$ with Jacobian matrix $\left[\dfrac{\partial \tilde{f}^j}{\partial x^i}\right]$ where $\tilde{f} = \mathbf{y} \circ f \circ \mathbf{x}^{-1} : \mathbf{x}(U \cap f^{-1}(V)) \to \mathbb{R}^{\dim N}$. Namely

$$df_p\Big(\frac{\partial}{\partial x^i}\Big|_p\Big) = \sum_j \frac{\partial \tilde{f}^j}{\partial x^i}(\mathbf{x}(p))\frac{\partial}{\partial y^j}\Big|_p.$$

Two special cases with one of the manifolds being $\mathbb{R}$ are particularly interesting.

Example 1.13. The first is the *total differential* df of a function $f : M \to \mathbb{R}$. Let y be the coordinate of $\mathbb{R}$, then $df_pX = a\partial/\partial y|_p$ for some a. By substituting $h = y$ in the definition of df_p we get $a = Xf$.

Since there is only one basis $\partial/\partial y|_p$ for $T_p\mathbb{R}$, following the usual convention we denote a vector in $\mathbb{R}$ simply by its coefficient. Also we drop the point p if no confusion is likely to occur. Hence

$$df(X) = Xf.$$

Each coordinate x^i is a C^∞ function at p and we compute

$$dx^i\Big(\frac{\partial}{\partial x^j}\Big) = \frac{\partial x^i}{\partial x^j} = \delta^i_j.$$

That is, the differentials dx^i's form a dual basis of the *cotangent space*, the dual space of tangent space:

$$T^*_pM := (T_pM)^* = \mathrm{Hom}(T_pM, \mathbb{R})$$

with respect to the basis $\partial/\partial x^i$'s. Moreover,

$$df = \sum_i \frac{\partial(f \circ \mathbf{x}^{-1})}{\partial x^i}\, dx^i.$$

This follows by looking at the values of both sides on the basis vectors $\partial/\partial x^i$'s.

Example 1.14. The second example is *the tangent vector*

$$\gamma'(t) = d\gamma_t\Big(\frac{\partial}{\partial t}\Big)$$

of a parameterized C^1 curve $\gamma : (a,b) \subset \mathbb{R} \to M$ with parameter $t \in (a,b)$. Here, following the usual convention, we identify $\gamma'(t) \equiv d\gamma_t$ as its image vector since there is one basis vector $\partial/\partial t$ on $T_t(a,b)$.

In local chart $(U, \mathbf{x})$, the curve is represented by

$$t \mapsto \tilde{\gamma}(t) := \mathbf{x} \circ \gamma(t) = (x^1(t), \ldots, x^m(t))^t$$

and

$$\gamma'(t) = \sum_i (x^i)'(t) \, \frac{\partial}{\partial x^i}\Big|_{\gamma(t)}.$$

The way df_p approximates f is best explained through the inverse/implicit function theorem. We start with the simplest notions, namely the injectivity and surjectivity of f.

Definition 1.15. A map $f \in C^\infty(M, N)$ is an *immersion* at $p \in M$ if the linear map df_p is injective, it is a *submersion* at p if df_p is surjective.

Lemma 1.16. *Let $f : M^m \to N^n$ be a C^∞ map.*

(1) *If f is an immersion at p (so $m \le n$), then there are charts $(U, \mathbf{x})$ at p and $(V, \mathbf{y})$ at $f(p)$ such that $f|_U$ is represented by $y^i = \tilde{f}^i(x) = x^i$ for $i = 1, \ldots, m$ and $y^i = 0$ for $i \ge m+1$. That is, U is a coordinate slice of V.*

(2) *If f is a submersion at p (so $m \ge n$), then there are charts $(U, \mathbf{x})$ at p and $(V, \mathbf{y})$ at $f(p)$ such that $f|_U$ is represented by $y^i = \tilde{f}^i(x) = x^i$ for $i = 1, \ldots, n$. That is, $\tilde{f}$ is a coordinate projection from U to V.*

PROOF. For (1), we start with any charts such that $f(U) \subset V$. Since $d\tilde{f}_{\mathbf{x}(p)}$ is injective, it has rank m. By reordering of coordinates y^i's', we may assume that the first $m \times m$ square matrix $\left[\dfrac{\partial \tilde{f}^i}{\partial x^j}(\mathbf{x}(p))\right]_{i,j=1}^m$ is invertible. Denote by

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 \end{bmatrix} = \tilde{f}(\mathbf{x}) = \begin{bmatrix} f^1(\mathbf{x}) \\ f^2(\mathbf{x}) \end{bmatrix}$$

under $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$, then by the inverse function theorem $\mathbf{y}_1 = f^1(\mathbf{x})$ is invertible over some $\mathbf{x}(p) \in W \subset \mathbf{x}(U)$. Using $(\mathbf{x}^{-1}(W), \mathbf{y}^1 =$

$f^1 \circ \mathbf{x})$ as a new chart at p and letting $g = f^2 \circ (f^1)^{-1}$, the map f becomes a graph of g:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^1 \\ g(\mathbf{y}^1) \end{bmatrix}.$$

By a simple change of coordinates

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}^1 \\ \mathbf{z}^2 \end{bmatrix} = \begin{bmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 - g(\mathbf{y}^1) \end{bmatrix}$$

near $f(p) \in N$, we get the desired coordinate charts.

For (2), again we start with any charts with $f(U) \subset V$. Since $d\tilde{f}_{\mathbf{x}(p)}$ is surjective, it has rank n. By reordering of coordinates x^i's we may assume that the first $n \times n$ square matrix $\left[\dfrac{\partial \tilde{f}^i}{\partial x^j}(\mathbf{x}(p))\right]_{i,j=1}^{n}$ is invertible. Denote by $\mathbf{x} = (\mathbf{x}^1, \mathbf{x}^2)^t$ under $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n}$ and consider the map $F : \mathbf{x}(U) \to \mathbf{y}(V) \times \mathbb{R}^{m-n}$ defined by

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{x}^2 \end{bmatrix} := F(\mathbf{x}^1, \mathbf{x}^2) = \begin{bmatrix} \tilde{f}(\mathbf{x}^1, \mathbf{x}^2) \\ \mathbf{x}^2 \end{bmatrix}.$$

Since

$$dF_{\mathbf{x}(p)} = \begin{bmatrix} D_1\tilde{f} & D_2\tilde{f} \\ 0 & \mathrm{id}_{m-n} \end{bmatrix}$$

is invertible, the inverse $G = F^{-1}$ exists over a smaller neighborhood $W \ni \mathbf{x}(p)$. The result follows by using $(\mathbf{y}, \mathbf{x}^2)^t$ as the new coordinate system at p. $\square$

Exercise 1.13. Show that $f \in C^\infty(M, N)$ can be locally represented by

$$\tilde{f}(\mathbf{x}) = (x^1, \ldots, x^k, 0, \ldots, 0)^t$$

for some $k \le m$ if and only if that df_p has constant rank k for all $p \in M$.

5. Submanifolds and the Whitney imbedding theorem

There is a well defined notion of sub-objects in a reasonably given category.

Definition 1.17. For a manifold N, a topological subspace $M \subset N$ is a *submanifold* if there is an atlas $\{(U_\alpha, \mathbf{x}_\alpha)\}_{\alpha \in A}$ on N such that the restriction

$$\{(U_\alpha \cap M, \mathbf{x}_\alpha|_{U_\alpha \cap M})\}_{\alpha \in A}$$

also form an atlas on M. This definition applies to any C^k manifolds.

Let $f : M \to N$ be an immersion. By lemma 1.16 (1), for any $p \in M$, there is a chart $U \ni p$ so that $f|_U$ is injective and $f(U)$ is a submanifold of N. However, f may not be injective globally, e.g. parameterized plane curves with self-intersections.

Even if f is an *injective immersion*, the image with the subspace topology might fail to be a manifold at all!

Example 1.18. Consider the plane curve in polar coordinates $r = \sin 2\theta$ with $\theta \in (0, \pi)$. The parametrization $\gamma : (0, \pi) \to \mathbb{R}^2$ given by

$$(x(\theta), y(\theta)) = (r \cos\theta, r \sin\theta) = (\sin 2\theta \cos\theta, \sin 2\theta \sin\theta)$$

is an injective immersion of $(0, \pi)$ into $\mathbb{R}^2$. But the point $(0, 0) \in \gamma((0, \pi))$ does not have any locally Euclidean neighborhood, when the image $\gamma((0, \pi))$ is equipped with the subspace topology in $\mathbb{R}^2$.

Even if the image is a manifold, it may not be equipped with the induced subspace topology:

Exercise 1.14. Let $a \in \mathbb{R} \setminus \mathbb{Q}$ and consider the map

$$f : \mathbb{R} \to S^1 \times S^1 : \qquad t \mapsto (e^{it}, e^{iat})$$

where we identify S^1 as a subset in $\mathbb{C}$. Show that f is an injective immersion and $f(\mathbb{R})$ is dense in $S^1 \times S^1$.

Definition 1.19. A C^∞ map $f : M \to N$ is an *imbedding* if it is an injective immersion which induces a homeomorphism $f : M \xrightarrow{\sim} f(M)$ with $f(M) \subset N$ being equipped with the subpace topology.

Lemma 1.20. *If $f : M \to N$ is a C^∞ imbedding then $f(M) \subset N$ is a C^∞ submanifold and $f : M \to f(M)$ is a diffeomorphism.*

PROOF. The condition that f being a homeomorphism means that for U a open neighborhood at p, there are open set $V \subset N$ at $f(p)$ such that $f^{-1}(V) = U$ and U is homeomorphic to $f(U) = V \cap f(M)$ under f. By lemma 1.16 (1) we may select U to be a coordinate slice of V and hence $f(M)$ is a C^∞ submanifold of N. $f : M \to f(M)$ is a diffeomorphism since they have identically the same atlas. □

A continuous map $f : M \to N$ between topological spaces is called *closed* if the image of a closed set is closed. It is clear that an injective continuous closed map induces a homeomorphism onto its image, so an injective closed immersion is an imbedding.

Similarly f is *open* if it send open sets to open sets. An injective open immersion is also an imbedding. However, an imbedding needs not be closed nor open. E.g. an interval (a, b) along the x-axis in $\mathbb{R}^2$.

A continuous map is *proper* if the inverse image of a compact set is compact.

Exercise 1.15. Let $f \in C^0(M, N)$ with N being Hausdorff. Show that: (1) If M is compact then f is proper as well as closed. (2) If M, N are manifolds and f is proper then it is also closed.

Thus for compact domain manifolds there is no serious topological issues to concern; the notion of immersions and imbeddings (= injective immersions here) are precise and convenient. For non-compact domain manifolds, extra information on f are usually crucial.

The Whitney imbedding theorem says that any manifold is nothing more than an imbedded submanifold in the Euclidean space. Before proving this fundamental result, we need a definition.

Definition 1.21. A set $A \subset \mathbb{R}^d$ has *measure zero* if for any $\epsilon > 0$ there is a countable cover by balls B_i with $\sum_i \mathrm{vol}(B_i) < \epsilon$. A set $A \subset M^d$ in a C^k manifold ($k \geq 1$) has measure zero if for any chart $(U, \mathbf{x})$ the set $\mathbf{x}(A \cap U)$ has measure zero in $\mathbb{R}^d$.

By the standard diagonal argument we see that a countable union of measure zero sets also has measure zero. Also it is clear that measure zero sets can not contain open sets. To see that the later definition makes sense we need "measure zero" to be independent of the choices of coordinates. Indeed, more is true:

Exercise 1.16. (1) If $f : U \to \mathbb{R}^d$ is C^1 and $A \subset U \subset \mathbb{R}^d$ has measure zero then $f(A)$ also has measure zero. (2) If $f : M^m \to N^n$ is C^1 and $m < n$, then $f(M)$ has measure zero, in particular it is not surjective.

Theorem 1.22 (Whitney Imbedding Theorem (weak form), 1936)**.** *Every C^∞ manifold M^d admits a C^∞ closed imbedding into $\mathbb{R}^{2d+1}$ and a C^∞ closed immersion in $\mathbb{R}^{2d}$.*

PROOF. We will only give the proof for the simpler case when M is assumed to be compact.

Step 1: Construct an imbedding $f : M \to \mathbb{R}^N$ for some large $N \in \mathbb{N}$. (This step requires only the C^1 structure.)

For any $p \in M$, consider a chart $(U_p, \mathbf{x}_p)$ with $\mathbf{x}_p(U_p) = B_0(2)$. The open cover $\{U'_p := \mathbf{x}_p^{-1}(B_0(1))\}_{p \in M}$ admits a finite subcover indexed by $1, \dots, k$. Consider cut off functions $\{\psi_i\}_{i=1}^k$ with $\psi_i \equiv 1$ on U'_i and $\operatorname{supp} \psi_i \subset U_i$. Define a C^∞ map

$$\begin{aligned} f &:= \prod_{i=1}^{k} (\psi_i \mathbf{x}_{p_i}, \psi_i) \\ &\equiv (\psi_1 \mathbf{x}_{p_1}, \psi_1, \cdots, \psi_k \mathbf{x}_{p_k}, \psi_k) : \quad M \to (\mathbb{R}^{d+1})^k = \mathbb{R}^{k(d+1)}. \end{aligned}$$

To see that f is an immersion, we notice that

$$df_p = (d(\psi_1 \mathbf{x}_{p_1}), d\psi_1, \cdots, d(\psi_k \mathbf{x}_{p_k}), d\psi_k) : T_pM \to T_{f(p)}\mathbb{R}^{k(d+1)}.$$

Let $p \in U'_i$. Since $\psi_i|_{U'_i} \equiv 1$, we get $d(\psi_i \mathbf{x}_{p_i})_p = d(\mathbf{x}_{p_i})_p$. This is the identification map $T_pM \cong T_{\mathbf{x}_{p_i}(p)}\mathbb{R}^d$ hence in particular that df_p is injective.

To see that f is injective, given $p \neq p' \in M$, if there is an i such that $p, p' \in U'_i$, then the component $\psi_i \mathbf{x}_{p_i} = \mathbf{x}_{p_i}$ gives different coordinates for p and p'. Otherwise $p \in U'_i$ and $p' \notin U'_i$ for some i and then $\psi_i(p) = 1 > \psi_i(p')$.

Step 2: Reduction of imbedding dimension N to $2d+1$.

This step works for any given closed imbedding $f : M \to \mathbb{R}^N$ (we need only the C^2 condition on M and f, and M may be non-compact). The idea is find a direction $v \in S^{N-1}$ and compose f with the projection map $\pi_v : \mathbb{R}^N \to v^\perp \cong \mathbb{R}^{N-1}$. The new map $f_v := \pi_v \circ f$ is closed since projection maps are clearly closed maps and composition of closed maps are again closed.

To have f_v being injective it is equivalent to require that for any $p \neq q \in M$, the vector $\overrightarrow{f(p)f(q)}$ is not parallel to v. More precisely, let $\Delta : M \to M \times M$ be the diagonal map $\Delta(p) = (p, p)$ and consider the *secant map*

$$\sigma : M \times M \backslash \Delta(M) \to S^{N-1}/\{\pm 1\} =: \mathbb{RP}^{N-1}$$

defined by

$$\sigma((p,q)) = \pm \frac{f(p) - f(q)}{\|f(p) - f(q)\|}.$$

Since σ is a C^∞ map from a $2d$ manifold to an $N-1$ manifold, if $2d < N-1$ (that is, $N > 2d+1$) then σ can not be surjective; in fact $\operatorname{Im}\sigma$ has measure zero. Thus f_v is injective if we select $v \notin \operatorname{im}\sigma$.

To have f_v being an immersion it is equivalent to require that $d(\pi_v)_x = \pi_v$ is injective on $T_x f(M)$ for all $x \in f(M)$. Without loss of generality we identify M as its image $f(M)$ in $\mathbb{R}^N$. Then the tangent bundle $TM := \bigcup_{x \in M} T_x M \subset M \times \mathbb{R}^N$ as a C^∞ manifold of dimension $2d$ (cf. the exercise below) and the unit sphere bundle

$$S(TM) \subset M \times S^{N-1}$$

as a C^∞ manifold of dimension $d + (d-1) = 2d-1$ is defined. The map

$$T : S(TM) \to S^{N-1}/\{\pm 1\} : \qquad (p, v) \mapsto \pm v$$

is a C^∞ map from a $2d-1$ manifold to an $N-1$ manifold. If $2d-1 < N-1$ (that is, $N > 2d$) then again $\operatorname{im} T$ has measure zero and T is not surjective. Thus f_v is an immersion if we select $v \notin \operatorname{im} T$.

When $N > 2d+1$, $\operatorname{im}\sigma \cup \operatorname{im} T \subset S^{N-1}$ also has measure zero hence the desired projection direction v can be selected. This completes the proof on the embedding statement. When $N = 2d+1$, we may still select the further projection f_v to get an immersion. This completes the proof. □

Exercise 1.17. For a C^k manifold M with $k \geq 1$, show that TM is a C^{k-1} manifold by constructing an atlas on it and computing the transition functions.

Show also that $S(TM)$ is a C^{k-1} submanifold of TM.

A similar idea leads to applications to homotopy theory:

Theorem 1.23. *If $f : S^k \to S^n$ is continuous with $k < n$, then f is homotopic to a constant map. That is, $\pi_k(S^n) = 0$ for $k < n$.*

PROOF. It is trivial if f is not onto since $S^n \setminus \{p\} \cong \mathbb{R}^n$ is contractible. This is indeed the case if f is C^1 by Exercise 1.16.

Now the idea is simply to approximate f by a C^1 (in fact C^∞) function $\tilde{f} : S^k \to S^n$ within a δ-error with $\delta < \pi$.

Exercise 1.18. Prove the C^∞ approximation for $f \in C(S^k, S^n)$ within any $\delta > 0$. In fact show that the C^1 approximation is always possible for any $f \in C(M, N)$ where M, N are both compact C^1 manifolds.

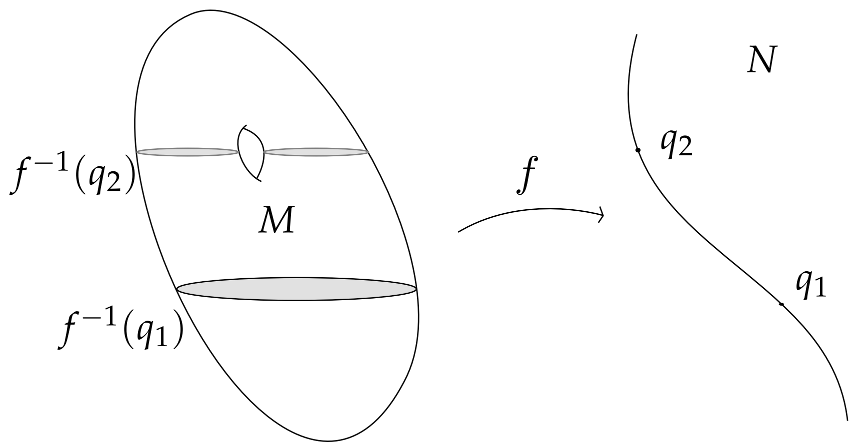
With this done, then for each $x \in S^k$, the two vectors $f(x)$, $\tilde{f}(x)$ span a two dimensional plane $V_x \subset \mathbb{R}^{n+1}$ and there is a unique homotopy $F(x,t)$ from $f(x)$ to $\tilde{f}(x)$ through the shorter great circle $V_x \cap S^n$. $F(x,t)$ is clearly continuous, hence f is homotopic to $\tilde{f}$, which is C^1 and hence homotopic to a constant map. □

Exercise 1.19 (Invariance of dimensionality). As a corollary, show that $\mathbb{R}^n \cong \mathbb{R}^m$ (homeomorphic) $\Longleftrightarrow n = m$.

Remark 1.24. Step 1 in our proof of Theorem 1.22 does not apply to the non-compact case. More transversality and approximation arguments are needed. Then, as in Step 2, Whitney produced an immersion $f : M^d \to \mathbb{R}^{2d}$ with at most transverse self-intersections (points). Finally, a surgery argument now called the Whitney trick removes the intersections and achieves the embedding $M^d \subset \mathbb{R}^{2d}$. This is known as the (strong form of) Whitney Imbedding Theorem and the bound is optimal. For details see e.g. [Hir94].

6. Submersions and Sard's theorem

After discussing submanifolds induced from immersions, we now consider those induced from submersions, i.e. df is surjective.



Definition 1.25. Let $f : M \to N$ be a smooth map. A point $q \in N$ is called a *regular value* of f if df_p is surjective for all $p \in f^{-1}(q)$. Otherwise q is called a *critical value* or *singular value*.

When q is a regular value, the preimage $f^{-1}(q)$ is a, possibly non-connected, submanifold of M and f can be parametrized as a coordinate projection locally at p by Lemma 1.16. Any point $p \in f^{-1}(q)$ is referred as a *regular point*.

A point $p \in M$ is called a *critical point* of f if p is not regular, i.e. df_p is not surjective. If $N = \mathbb{R}$, this means $df_p = 0$. We denote by $C(f)$ the set of all critical points.

Intuitively, the map f establishes a kind of nice *local fiber space structure* on M outside the singular values $f(C(f))$. Thus it is important to know more about the properties of $f(C(f))$.

Theorem 1.26 (Sard's Theorem, 1942)**.** *$f(C(f))$ has measure 0 in N.*

PROOF. Since a countable union of measure zero sets is measure zero, it suffices to prove the case of charts $(U, \mathbf{x})$ with bounded U.

Consider $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$. Let

$$C_i = \{\mathbf{x} \in U \mid D^\alpha f(\mathbf{x}) = 0,\ \forall\, \alpha,\ |\alpha| \le i\}$$

and $C = C(f)$. So $C \supset C_1 \supset C_2 \supset \cdots$.

This proof consists of three steps:

(1) $f(C \setminus C_1)$ has measure 0.
(2) $f(C_i \setminus C_{i+1})$ has measure 0.
(3) $f(C_k)$ has measure 0 for some k large enough.

We will prove them using induction on the total dimension $m+n$.

If $m+n = 1$, $C = C_1$ and $f(C)$ consist of only one point. The theorem is trivial in this case. So we assume that $m+n \geq 2$.

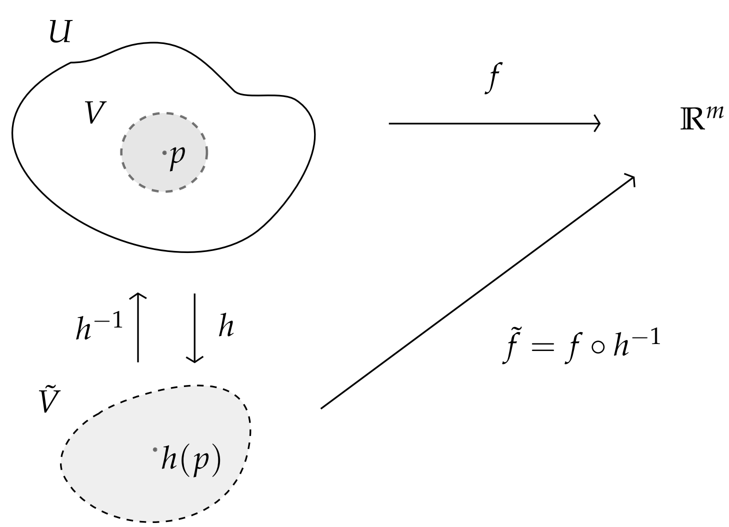
Let $p \in C \setminus C_1$, say $\partial f^1(p)/\partial x^1 \neq 0$. Through a coordinate change

$$\begin{array}{rccl} h: & U & \to & \mathbb{R}^m, \\ & x & \mapsto & (f^1(\mathbf{x}), x^2, \ldots, x^m)^t, \end{array}$$

we have

$$dh_p = \begin{bmatrix} \partial_1 f^1(p) & * \\ 0 & I_{m-1} \end{bmatrix}$$

which has a non-zero Jacobian. By the Inverse Function Theorem, there is a neighborhood $V \subset U$ such that h^{-1} exists on $\tilde{V} := h(V)$.



Let $\tilde{f} = f \circ h^{-1} : \tilde{V} \to \mathbb{R}^m$. We have $\tilde{f}^1(t, \ldots) = t$ and

$$q \in C(\tilde{f}) \cap \tilde{V} \iff h^{-1}(q) \in C(f) \cap V.$$

Then $\tilde{f} : \tilde{V} \to \mathbb{R}^n$ splits into

$$\tilde{f}_t : (\{t\} \times \mathbb{R}^{m-1}) \cap \tilde{V} \to \{t\} \times \mathbb{R}^{n-1},$$

and

$$d\tilde{f} = \begin{bmatrix} 1 & 0 \\ * & df_t \end{bmatrix}, \quad q = (t, r) \in C(\tilde{f}) \iff r \in C(\tilde{f}_t).$$

By the induction hypothesis, $\tilde{f}_t(C(\tilde{f}_t))$ has measure zero in the hyperplane $\{t\} \times \mathbb{R}^{n-1}$. Since $C(\tilde{f}) \setminus C_1 \subset \bigcup_t C(\tilde{f}_t)$, by Fubini's theorem in the theory of Lebesgue integrals,

$$|f(C(\tilde{f}) \setminus C_1)| \leq \int_t \left|\tilde{f}_t(C(\tilde{f}_t))\right| dt = 0.$$

Remark 1.27. This argument does not really need the full power of Lebesgue theory. We only need the theory of measure 0 for the proof.

Consider the same argument on all $p \in C \setminus C_1$. Since $C \setminus C_1$ can be covered by countable union of such $(C(\tilde{f}) \setminus C_1)$'s, we conclude that $f(C \setminus C_1)$ has measure 0.

Secondly, for any $p \in C_i \setminus C_{i+1}$, we may assume $D^\alpha f(p) = 0$ for all $|\alpha| \leq i$ but $D^\beta f^1(p) \neq 0$ for some $\beta = \alpha + (1, 0, \ldots, 0)$.

Write $f^{(\alpha)}(\mathbf{x}) = D^\alpha f^1(\mathbf{x})$. Again by changing coordinates,

$$\begin{aligned} h : \ U &\to \mathbb{R}^m, \\ x &\mapsto (f^\alpha(\mathbf{x}), x^2, \ldots, x^m)^t, \end{aligned}$$

$$dh_p = \begin{bmatrix} \partial_1 f^{(\alpha)}(p) & * \\ 0 & I_{m-1} \end{bmatrix}$$

is invertible and there is a neighborhood $V \subset U$ such that h^{-1} exists. Also, $h(C_i \cap V) \subset \{0\} \times \mathbb{R}^{m-1}$. Let $\tilde{f} = f \circ h^{-1}$. We have

$$h^{-1}(q) \in C_i(f) \iff q \in C_i(\tilde{f}) = C_i(\tilde{f}_0)$$

where $\tilde{f}_0 : (\{0\} \times \mathbb{R}^{m-1}) \bigcap h(V) \to \mathbb{R}^n$.

By induction, $f(C(\tilde{f}_0))$ has measure 0, and so does $f(C_i \setminus C_{i+1})$ by the countable covering argument as before.

Thirdly, we claim that $f(C_k)$ has measure 0 for large k through a volume estimate. By the Taylor expansion, there exists a constant A which depends only on k, m and n such that for $p \in C_k$,

$$f(p+h) = f(p) + R(p,h), \quad |R(p,h)| \le A|h|^{k+1}$$

for all $|h| < \delta$. Let $v_n = \mathrm{vol}(B_0(1))$. In each such ball $B_p(\delta)$ we have

$$|f(C_k \cap B_p(\delta))| \le v_n A^n \delta^{(k+1)n}$$

Now we cover C_k by a finite number of such balls. We need at most $(2d/\delta)^m$ balls with $d = \operatorname{diam} U$. Pick k satisfying $k+1 > m/n$, then

$$|f(C_k)| \le v_n A^n (2d)^m \delta^{(k+1)n-m}.$$

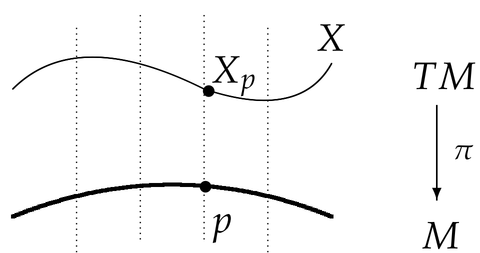
Since A and d are independent of δ, by taking $\delta \to 0$ we get that $f(C_k)$ has measure 0. The proof is complete. □

Remark 1.28. Sard's Theorem holds for C^k maps with $k > \max\{m-n, 0\}$. The case $n=1$ was first proved by Morse in 1939.

7. Vector fields, flows, Lie derivatives and the Frobenius integrability theorem

Definition 1.29. Let $U \subset M$ be open. A map $X : U \to TM$ is called a *vector field* if for each $p \in U$, $X(p) \in T_pU = T_pM$. That is, X is a section of the tangent bundle over U.

We say that X is a C^k and write $X \in C^k(U, TM)$ if $Xf \in C^k(U)$ for all $f \in C^k(U)$. The space of C^∞ (smooth) vector fields on U are denoted by $T(U) := C^\infty(U, TM)$.



We have seen that TM is a smooth manifold where charts and coordinate functions on TM are of the form

$$(\pi^{-1}(U), x^1, \cdots, x^m, dx^1, \cdots, dx^m).$$

So we also have equivalent definitions for C^∞ vector fields:

Proposition 1.30. *The followings are equivalent:*

(1) $X(f) \in C^\infty(U)$ *for any* $f \in C^\infty(U)$*;*

(2) $X = \sum a^i(\partial/\partial x^i)$*,* $a^i \in C^\infty(U)$*;*

(3) $X : U \to TM$ *is* C^∞*.*

In Euclidean spaces, for a Lipschitz vector field $F : U \subset \mathbb{R}^m \to \mathbb{R}^m$ we can find an unique solution to the initial value problem

$$\begin{cases} \dfrac{\partial \mathbf{x}}{\partial t} = F(\mathbf{x}(t)), \quad \mathbf{x} = (x^1(t), \dots, x^m(t)) \\ \mathbf{x}(0, \mathbf{x}_0) = \mathbf{x}_0 \end{cases}$$

which depends continuously on $\mathbf{x}_0$. This is known as the *existence and uniqueness theorems of ODE* [Picard 1890, Lindelöf 1894] in basic ODE courses. An improved version to take care of the smooth dependence on the initial conditions is as follows:

Theorem 1.31. *For a* C^k *vector field* F *with* $k \geq 1$*, there exists a unique solution* $\mathbf{x}(t, \mathbf{x}_0)$ *which is* C^{k+1} *in* t *and* C^k *in* $(t, \mathbf{x}_0)$*.*

We will apply it in the case $k = \infty$ in this section. A proof of the improved statement will be given in the next section.

There are corresponding versions on manifolds. Firstly we have:

Theorem 1.32. *Let* X *be a* C^∞ *vector field on* M*. For any* $p \in M$*, there exists* $(a_p, b_p) \subset \mathbb{R} \bigcup \{\pm\infty\}$ *and* $\gamma_p : (a_p, b_p) \to M$ *such that*

$$\gamma_p(0) = p, \qquad \gamma_p'(t) = X(\gamma_p(t))$$

and γ_p *is maximal among curves with this property.*

The solution γ_p is called an *integral curve* of X. Furthermore, for any fixed $t \in \mathbb{R}$, it is useful to consider the *flow generated by* X

$$\phi_t(p) := \gamma_p(t)$$

on the domain

$$\mathscr{D}_t := \{p \in M \mid t \in (a_p, b_p)\}.$$

Clearly, ϕ_t is a diffeomorphism from $\mathscr{D}_t$ to $\mathscr{D}_{-t}$.[1] And $\{\phi_t\}_{t\in\mathbb{R}}$ defines a one-parameter group of (local) diffeomorphisms:

- $\phi_t \circ \phi_s = \phi_{t+s}$
- $\phi_t^{-1} = \phi_{-t}$

It is fundamental and natural to ask: *how to take derivatives of a vector field?* Notice that it does not make sense to compare the two vectors X_p and X_q when $p \neq q$. In $\mathbb{R}^d$, we take derivatives by moving $f(\gamma(t))$ back to $\gamma(0)$ and calculate

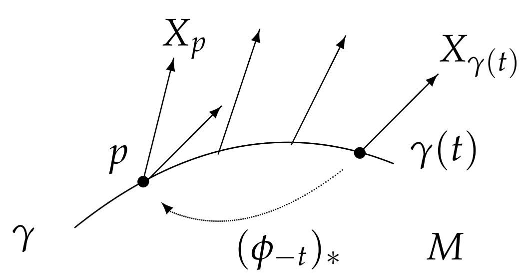
$$v_p(f) = \lim_{t\to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t}$$

along any C^1 curve γ with $\gamma(0) = p$ and $\gamma'(0) = v$.

On a manifold, one solution is the *Lie derivative*.

Definition 1.33. Given two vector fields V, X. Let ϕ_t be the flow generated by V. The Lie derivative of X along V is defined by

$$\begin{aligned} L_V X(p) &:= \lim_{t\to 0} \frac{1}{t}\left((\phi_{-t})_* X_{\phi_t(p)} - X_p \right) \\ &= \frac{d}{dt}\Big|_{t=0} (\phi_{-t})_* X_{\phi_t(p)}. \end{aligned}$$



We shall see shortly that the Lie derivative corresponds to another simple binary operation on vector fields, the *Lie bracket*.

---

[1] We remark that if M is compact, the interval (a_p, b_p) can be extended to $(-\infty, \infty)$ and ϕ_t will be a diffeomorphism from M to itself. In general, a vector field is called **complete** if it has this property.

Definition 1.34. Let X, Y be C^∞ vector fields on U. The Lie bracket is defined by $[X, Y] := XY - YX$. That is,

$$[X, Y]_p f = X_p(Y(f)) - Y_p(X(f)).$$

It seems that $[X, Y]$ is a second order operator on $C^\infty(U)$. In fact, it is still of the first order:

Proposition 1.35. *$[X, Y]$ is a C^∞ vector field on U.*

PROOF. Let $X = \sum_i a^i \frac{\partial}{\partial x^i}$, $Y = \sum_j b^j \frac{\partial}{\partial x^j}$ and $f \in C^\infty$. By Leibniz rule,

$$\begin{aligned}
[X, Y]_p f &= \sum_{ij} a^i \frac{\partial}{\partial x^i}\left(b^j \frac{\partial}{\partial x^j} f\right) - b^j \frac{\partial}{\partial x^j}\left(a^i \frac{\partial}{\partial x^i} f\right) \\
&= \sum_{ij} a^i \frac{\partial b^j}{\partial x^i}\frac{\partial f}{\partial x^j} + a^i b^j \frac{\partial}{\partial x^i}\frac{\partial f}{\partial x^j} - b^j \frac{\partial a^i}{\partial x^j}\frac{\partial f}{\partial x^i} - b^j a^i \frac{\partial}{\partial x^j}\frac{\partial f}{\partial x^i} \\
&= \sum_{ij}\left(a^i \frac{\partial b^j}{\partial x^i}\frac{\partial}{\partial x^j} - b^j \frac{\partial a^i}{\partial x^j}\frac{\partial}{\partial x^i}\right) f.
\end{aligned}$$

□

The Lie bracket has the following basic (Lie algebra) properties:

Proposition 1.36. *Let X, Y be C^∞ vector fields. Then*

(1) *(Anti-symmetry)* $[X, Y] = -[Y, X]$
(2) *(Jacobi identity)* $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$

Now we can prove the important formula:

Theorem 1.37 (Lie, 1870)**.** $(L_V W)_p = [V, W]_p$

PROOF. Let V generate ϕ_t and W generate ψ_s. Choose a test function $h \in C_p^\infty$. By definition,

$$
\begin{aligned}
&(L_V W)_p\, h \\
&= \frac{d}{dt}\Big|_{t=0} (\phi_{-t})_* W_{\phi_t(p)} h = \frac{d}{dt}\Big|_{t=0} W_{\phi_t(p)} h \circ \phi_{-t} \\
&= \frac{d}{dt}\frac{d}{ds}\, h \circ \phi_{-t} \circ \psi_s(\phi_t(p))\Big|_{t=0,s=0} \\
&= \frac{d}{ds}\frac{d}{dt}\, h \circ \phi_{-t} \circ \psi_s \circ \phi_t(p)\Big|_{t=0,s=0} \\
&= \frac{d}{ds}\left(dh\left(-V_{\phi_{-t}\circ\psi_s\circ\phi_t(p)}\right) + dh \circ {\phi_{-t}}_* \circ {\psi_s}_* \left(V_{\phi_t(p)}\right)\right)\Big|_{t=0,s=0} \\
&= \frac{d}{ds}\left(-dh(V_{\psi_s(p)}) + d(h\circ\psi_s)(V_p)\right)\Big|_{s=0} \\
&= -\frac{d}{ds}\Big|_{s=0} Vh(\psi_s(p)) + V_p\left(\frac{d}{ds}\Big|_{s=0} h\circ\psi_s(p)\right) \\
&= -W_p(Vh) + V_p(Wh) = [V,W]_p\, h.
\end{aligned}
$$

□

We may extend these concepts to some higher dimensional cases.

Definition 1.38. A k dimensional *distribution* $\mathscr{D}$ on a manifold M is a choice of a k dimensional subspace $\mathscr{D}(p)$ of T_pM for each $p \in M$.

We say $\mathscr{D}$ is C^∞ if for every $p \in M$, there is a neighborhood U such that $\mathscr{D}$ is spanned by k C^∞ vector fields $X_1, X_2, \ldots, X_k$ on U.

The question in higher dimensions becomes: for a given distribution $\mathscr{D}$ near p, does there exists a k dimensional submanifold S with $p \in S$ such that $T_qS = \mathscr{D}(q)$ for all $q \in S$?

If so, S is called an integral manifold of $\mathscr{D}$ passing through p. Clearly a necessary condition for this integrability is the following:

Definition 1.39. $\mathscr{D}$ is called *involutive* if $[X,Y] \in \mathscr{D}$ for all $X, Y \in \mathscr{D}$. Here a vector field $X \in \mathscr{D}$ means $X_p \in \mathscr{D}(p)$ for all p.

This turns out to be sufficient at least locally:

Theorem 1.40 (Frobenius Integrability, 1877)**.** *If $\mathscr{D}$ is involutive, then for all $p \in M$ there exists a maximal integral manifold $S_p \subset M$ passing through p such that $T_qS_p = \mathscr{D}(q)$ for all $q \in S_p$.*

As in the case of integral curves, $S_p \subset M$ needs not be closed. That is S_p may not be with the induced topology.

The proof is based on two lemmas. To state the first one, we need

Definition 1.41. For a smooth map $f : S \to M$ and a smooth vector field $X \in C^\infty(TS)$, the tangent map $df = f_*$ sends X to

$$f_* X \in C^\infty(S, f^*TM),$$

namely a section of the pull back tangent bundle. Here $(f^*TM)_p := T_{f(p)}M$. If f is not injective, say $f(p) = f(q)$, it might happens that $f_{*,p}X_p \neq f_{*,q}X_q$, hence f_*X might not be the restriction of a vector field X' on M.

If indeed $f_*X = X'|_{f(S)}$ for some $X' \in C^\infty(TM)$, we say that X is a *f-related* vector field (with X'), or X and X' are f-related.

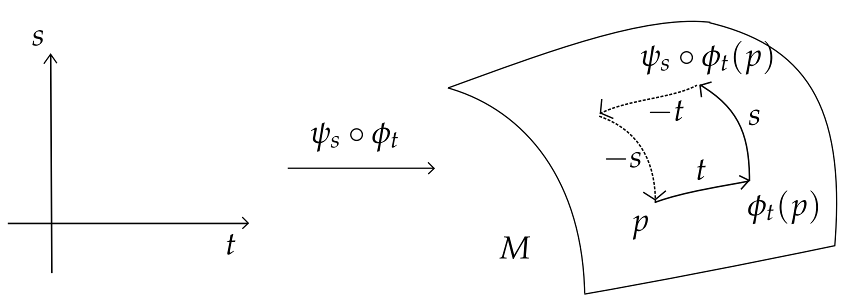
Exercise 1.20. Let $f \in C^\infty(S, M)$. Show that $[f_*X, f_*Y]$ is defined for f-related fields $X, Y \in C^\infty(TS)$ and $f_*([X, Y]) = [f_*X, f_*Y]$.

The next lemma characterizes coordinate vector fields:

Lemma 1.42. *If $X_1, \ldots, X_k$ are k vector fields near a point $p \in M$ that spans $\mathscr{D}$ and $[X_i, X_j] = 0$ for each i, j, then locally near p there exists a coordinate submanifold $S \ni p$ such that $X_i = \partial/\partial x^i$, $i = 1, 2, \cdots, k$.*

That is, there exists a coordinate system $(U, \mathbf{x})$ at p such that S can be locally represented as $\{x^j = 0 \mid j = k+1, \cdots, m\}$

PROOF. Assume $[V, W] = 0$. Let V generate ϕ and W generate ψ. We claim that $\phi_t \circ \psi_s(p) = \psi_s \circ \phi_t(p)$.



Let $c(t,s) = \psi_{-s} \circ \phi_{-t} \circ \psi_s \circ \phi_t(p)$. Since $L_V W = 0$ implies $\phi_{-t*}W = W$, directly differentiation shows that

$$\begin{aligned}\frac{\partial}{\partial s}c(t,s) &= -W_{\psi_{-s}\circ\phi_{-t}\circ\psi_s\circ\phi_t(p)} + \psi_{-s*}\circ\phi_{-t*}(W_{\psi_s\circ\phi_t(p)}) \\ &= -W_{c(t,s)} + W_{c(t,s)} = 0\end{aligned}$$

Similarly, $\partial c(t,s)/\partial t = 0$. So $c(t,s) = p$ for all well-defined t, s.

Now, suppose X_i generates ϕ^i for $i = 1, 2, \ldots, k$ near p. We define

$$c(t^1, t^2, \ldots, t^k) := \phi^1_{t^1} \circ \phi^2_{t^2} \circ \cdots \circ \phi^k_{t^k}.$$

We check that its coordinate tangent vectors are precisely $X_1, \ldots, X_k$:

$$c_*(\partial/\partial t^1) = \frac{\partial}{\partial t^1}\phi^1_{t^1} \circ \cdots \circ \phi^k_{t^k} = X_1\Big|_{c(t^1,\ldots,t^k)}.$$

And since every $\phi^i_{t_i}$ commutes with each other,

$$c_*(\partial/\partial t^i) = \frac{\partial}{\partial t^i}\phi^i_{t^i} \circ \phi^1_{t^1} \circ \cdots \widehat{\phi^i_{t^i}} \cdots \circ \phi^k_{t^k} = X_i\Big|_{c(t^1,\ldots,t^k)}.$$

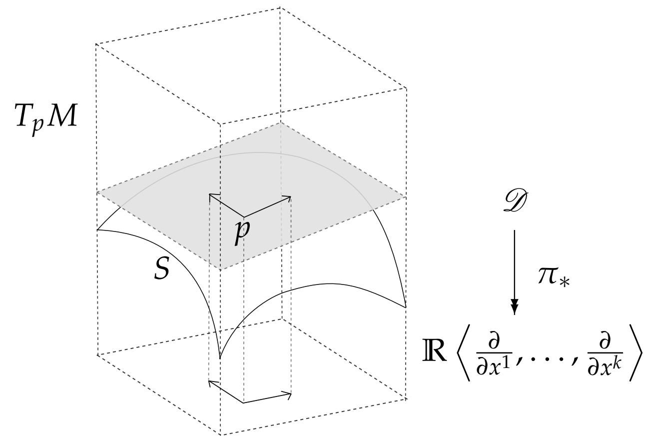
Let $S := \{c(t^1, \ldots, t^k) \mid t^j \in I_j, \forall\, j\}$ where I_j is chosen to be small enough such that every $\phi^j_{t^j}$ is well-defined for $t^j \in I_j$. We see that c_* is injective and then c is an immersion into M. By Lemma 1.16, there exists a chart $(U, \mathbf{x})$ near p such that $x^i(c(t^1, \ldots, t^k)) = t^i$ for $i = 1, \ldots, k$ and $x^i(c(t^1, \ldots, t^k)) = 0$ for $i = k+1, \ldots, m$. Hence $S \cap U$ is a coordinate slice $\{x^j = 0 \mid j = k+1, \ldots, m\}$ on M. □

PROOF OF FROBENIUS THEOREM. We fix an arbitrary point $p \in M$. Since $\mathscr{D}(p) \subset T_pM$ is a k dimensional subspace, we can select a chart $(U, \mathbf{x})$ at p with $\mathscr{D}(p) = \mathbb{R}\left\langle \partial_1|_p, \ldots, \partial_k|_p \right\rangle$ and construct the projection $\pi : U \subset M \to \mathbb{R}^k$ onto the first k coordinates. Then there exists a smaller neighborhood $U' \ni p$ such that

$$\mathscr{D} \cong \pi_*\mathscr{D} = \mathbb{R}\left\langle \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^k} \right\rangle, \qquad \forall\, q \in U'.$$

For each $i = 1, \ldots, k$, let $X_i \in \mathscr{D}$ be the vector field lifted from $\partial/\partial x^i$, i.e. $\pi_*(X_i) = \partial/\partial x^i$. Then they are π-related and

$$0 = \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = [\pi_*(X_i), \pi_*(X_j)] = \pi_*[X_i, X_j] \Longleftrightarrow [X_i, X_j] = 0.$$



The above lemma then implies that there exists an integral manifold S passing through p such that $T_q S = \mathscr{D}(q)$ for all $q \in S$.

Since $p \in M$ is arbitrary, the union of all integral manifolds is the whole manifold M. Also for any two integral manifolds S and S', if $S \cap S' \neq \emptyset$ then $S \cup S'$ is also an integral manifold. We conclude that there is a maximal integral manifold S_p passing through p. $\square$

8. Existence, uniqueness and smooth dependence of ODE

Now we go back to the ODE system:

$$\begin{cases} X'(t) = F(X(t)), \quad F \in C^1(O, \mathbb{R}^n) \text{ a vector field on } O \subset \mathbb{R}^n, \\ X(0) = \mathbf{x}_0, \end{cases}$$

and assume the Picard–Lindelöf theorem that there exists a unique continuous function $\phi(t, \mathbf{x}_0)$ satisfying the equation for $t \in J$, a maximal interval for the existence of solutions, and with $\phi(0, \mathbf{x}_0) = \mathbf{x}_0$.[2]

It is obviously that $\frac{\partial}{\partial t}\phi(t, \mathbf{x}_0)$ exists. So, our goal is to discuss: *the smooth dependence of the solution $\phi(t, \mathbf{x}_0)$ on its initial value $\mathbf{x}_0$.*

It turns out that $\phi(t, \mathbf{x}_0)$ is C^1 in $\mathbf{x}_0$. Moreover, an iterative argument then implies the C^k case as stated in Theorem 1.32.

[2]In this section we work on $\mathbb{R}^n$ entirely and the symbols X, Y, Z etc. will be used to denote points in $\mathbb{R}^n$. This should not be confused with the same symbols in the last section which denote vector fields on a manifold.

Historically there exists two different proofs of this theorem, one goes through a classical method by estimates (c.f. [HSD13]) and the other makes use of the inverse function theorem on Banach spaces. Below we follow the first method closely.

Suppose there are two solutions $X(t)$ and $\tilde{X}(t)$ with the given initial data $\mathbf{x}_0$ and $\mathbf{x}_0 + \mathbf{z}_0$. The key point is to estimate $\|X(t) - \tilde{X}(t)\|$ in terms of $\mathbf{x}_0$ and $\mathbf{z}_0$. We consider the variational equation:

$$(*) \qquad \begin{cases} U'(t) = A(t)U(t), \\ U(0) = \mathbf{z}_0 \end{cases}$$

where $A(t) = F'(X(t))$, which is C^0 dependent on t. The idea is that, when $\mathbf{z}_0$ is small, $X(t) + U(t)$ should approximate $\tilde{X}(t)$ with initial data $\mathbf{x}_0 + \mathbf{z}_0$. In fact, this comes from the intuition that if F is C^2, then the solution to the variational problem is just the first order term of the Taylor expansion for $\tilde{X}(t)$'s in $\mathbf{z}_0$.

Proposition 1.43. *Let $U(t, \xi)$ be the flow of (*), i.e. $U(0, \xi) = \xi$, $\mathbf{x}_0 + \xi \in O$, and $Y(t, \xi)$ be the flow of $X'(t) = F(X(t))$ with $Y(0, \xi) = \xi$. Then*

$$\lim_{\|\xi\| \to 0} \frac{\|Y(t, \xi) - X(t) - U(t, \xi)\|}{\|\xi\|} = 0$$

uniformly on an interval of existence J.

Assuming the proposition, the theorem follows immediately.

Theorem 1.44. *If $F \in C^k$, then the flow $\phi(t, \mathbf{x}_0)$ of the ODE system $X'(t) = F(X(t))$, $X(0) = \mathbf{x}_0$ is C^k as well.*

PROOF. By the proposition,

$$\phi(t, \mathbf{x}_0 + \xi) - \phi(t, \mathbf{x}_0) = Y(t, \xi) - X(t) = U(t, \xi) + o(|\xi|).$$

Note that from solving the linear system $U'(t, \xi) = A(t)U(t, \xi)$ with $U(0, \xi) = \xi$, we see that

$$U(t, \xi) = e^{A(t)}\xi$$

is linear in ξ. Hence $D_2\phi(t, \mathbf{x}_0)\xi = U(t, \xi)$ and $\phi(t, \mathbf{x})$ is C^1 in $\mathbf{x}$. This proves the theorem for the case $k = 1$.

Back to the variational equation (*), we get

$$\frac{d}{dt}(D_2\phi(t,\mathbf{x}_0)) = F'(\phi(t,\mathbf{x}_0))D_2\phi(t,\mathbf{x}_0)$$

with $D_2\phi(0,\mathbf{x}_0) = \mathrm{id}_{\mathbb{R}^n}$. Then by induction, $F \in C^k$ implies that $\phi(t,\mathbf{x})$ is C^k in $\mathbf{x}$. □

PROOF OF PROPOSITION. We rewrite the differential equations into integral equations as:

$$\begin{aligned} X(t) &= \mathbf{x}_0 + \int_0^t F(X(s))ds, \\ Y(t,\xi) &= \mathbf{x}_0 + \xi + \int_0^t F(Y(s,\xi))ds, \\ U(t,\xi) &= \xi + \int_0^t F'(X(s))U(s,\xi)ds. \end{aligned}$$

By the Taylor expansion, we have an estimate:

$$\begin{aligned} &\|Y(t,\xi) - X(t) - U(t,\xi)\| \\ &\leq \int_0^t \|F(Y(s,\xi)) - F(X(s)) - F'(X(s))U(s,\xi)\|ds \\ &\leq \int_0^t \Big(|F'(X(s))\|\|Y(s,\xi) - X(s) - U(s,\xi)\| \\ &\qquad + \|R(X(s), Y(s,\xi) - X(s))\|\Big)ds \end{aligned}$$

where R is the first order remainder term.

We use the Gronwall's inequality to deal with the iteration of difference appearing in the integral.

Exercise 1.21 (Gronwall's inequality, an easy version). If $u \in C^1[0,d]$, $u > 0$ and u satisfies $u(t) \leq c + \int_0^t Ku(s)ds$ for some positive constants c, K. Then $u \leq ce^{Kt}$ on $[0,d]$.

Therefore, the constant $c = |\xi|$ can be taken to be small and

$$\begin{aligned} \|Y(t,\xi) - X(t)\| &\leq \|\xi\| + \int_0^t \|F(Y(s,\xi)) - F(X(s))\|ds \\ &\leq \|\xi\| + \int_0^t \|F'\|\|Y(s,\xi) - X(s)\|ds \end{aligned}$$

Choose K large such that $\|F'\| < K$ on a small neighborhood. By Gronwall's inequality, $\|Y(t,\xi) - X(t)\| \leq \|\xi\| e^{Kt}$. So, for any $\epsilon > 0$, we can choose ξ small such that the remainder term

$$\|R(X(s), Y(s,\xi) - X(s))\| \leq \epsilon \|Y(s,\xi) - X(s)\|.$$

Denote $g(t) = \|Y(t,\xi) - X(t) - U(t,\xi)\|$ and rewrite

$$g(t) \leq \int_0^t Kg(s) + \epsilon\|\xi\| e^{Ks} ds \leq \epsilon\|\xi\| C + \int_0^t Kg(s) ds$$

for some bounded constant C depending on F and the existece interval J. By Gronwall's inequality again, $g(t) \leq \epsilon\|\xi\| C e^{Kt}$ and hence

$$\frac{\|Y(t,\xi) - X(t) - U(t,\xi)\|}{\|\xi\|} \leq \epsilon \tilde{C}$$

which is uniformly in t. □

9. Problems

1. ([War83] Ch.1 #10) Let M be a compact manifold of dimension n, and let $f : M \to \mathbb{R}^n$ be C^∞. Prove that f cannot everywhere be non-singular.
2. ([War83] Ch.1 #3) Let $\{U_\alpha\}$ be an open cover of a manifold M. Prove that there exists a refinement $\{V_\alpha\}$ such that $\overline{V_\alpha} \subset U_\alpha$ for each α.
3. ([War83] Ch.1 #9) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by
$$f(x,y) = x^3 + xy + y^3 + 1.$$
For which points $p = (0,0)$, $p = (\frac{1}{3}, \frac{1}{3})$, $p = (-\frac{1}{3}, -\frac{1}{3})$ is $f^{-1}(f(p))$ an imbedded submanifold in $\mathbb{R}$?
4. ([War83] Ch.1 #16) Let $N \subset M$ be a submanifold. Let $\gamma : (a,b) \to M$ be a C^∞ curve such that $\gamma(a,b) \subset N$. Show that it is not necessarily true that $\dot{\gamma}(t) \in N_{\gamma(t)}$ for each $t \in (a,b)$.
5. ([War83] Ch.1 #17) Prove that any C^∞ vector field on a compact manifold is complete.
6. ([War83] Ch.1 #18) Prove that a C^∞ map $f : \mathbb{R}^2 \to \mathbb{R}^1$ cannot be one-to-one.
7. ([War83] Ch.1 #23) A Riemannian structure on a differentiable manifold M is a smooth choice of a positive definite inner product $\langle \, , \, \rangle_m$ on each tangent space M_m, smooth in the sense that whenever X and Y are C^∞ vector fields on M, then $\langle X, Y \rangle$ is a C^∞ function on M. Prove that there exists a Riemannian structure on every differentiable manifold. You will

need to use a partition of unity argument. A *Riemannian manifold* is a differentiable manifold together with a Riemannian structure.

8. ([War83] Ch.1 #6) Prove that if $\psi : M \to N$ is C^∞, one-to-one, onto, and everywhere non-singular, then ψ is a diffeomorphism.
9. ([War83] Ch.1 #19) Supply the details of the equivalence of the Frobenius theorem 1.40 and the classical version:

   Remark 1.45 (classical Frobenius theorem). Let U and V be open sets in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. We use coordinates $r_1, \dots, r_m$ on $\mathbb{R}^m$ and $s_1, \dots, s_n$ on $\mathbb{R}^n$. Let

   $$b : U \times V \to M(n,m)$$

   be a C^∞ map of $U \times V$ into the set of all $n \times m$ real matrices, and let $(r_0, s_0) \in U \times V$. If

   $$\frac{\partial b_{i\beta}}{\partial r_\gamma} - \frac{\partial b_{i\gamma}}{\partial r_\beta} + \sum_{j=1}^{n} \left(\frac{\partial b_{i\beta}}{\partial s_j} b_{j\gamma} - \frac{\partial b_{i\gamma}}{\partial s_j} b_{j\beta} \right) = 0$$

   $$(i = 1, \dots, n; \gamma, \beta = 1, \dots, m)$$

   on $U \times V$, then there exist neighborhoods U_0 of r_0 in U and V_0 of s_0 in V and a unique C^∞ map

   $$\alpha : U_0 \times V_0 \to V$$

   such that if

   $$\alpha_s(r) = \alpha(r,s) \quad (s \in V_0, r \in U_0)$$

   then

   $$\alpha_s(r_0) = s, \; d\alpha_s\big|_r = b(r, \alpha(r,s))$$

   for all $(r,s) \in U_0 \times V_0$.

10. ([War83] Ch.1 #20) Let $\varphi : N \to M$ be C^∞, and let X be a C^∞ vector field on N. Suppose that $d\varphi(X(p)) = d\varphi(X(q))$ whenever $\varphi(p) = \varphi(q)$. Is there a smooth vector field Y on M which is φ-related to X?