Analysis 3H

typed up by B. S. H. Mithrandir

- *Last compiled: July 2022*
- Adapted from notes of D. Schütz, Durham
- This was part of the Analysis 3H module elective. This is a course on real analysis, touching on metric spaces, tangent spaces, vector fields, manifolds, and differential forms.

- **TODO!** diagrams, notation (bold vs not bold)

1 Metric spaces

1.1 Basic notions

The field of real numbers $\mathbb{R}$ is a totally ordered field which also satisfies the **completeness** axiom, i.e. a non-empty bounded set $A \subseteq \mathbb{R}$ has a **supremum** and/or an **infimum**. The supremum of $A \subseteq \mathbb{R}$ is a real number s where $a \leq s$ for all $a \in A$. If m is also such that $a \leq m$ for $a \in A$, then $s \leq m$, denoted $\sup A$. The infimum of A is where the inequalities signs are swapped, denoted $\inf A$.

Lemma 1.1.1 *Let $I_n = [a_n, b_n]$ be a sequence of closed intervals such that $a_n \leq a_{n+1} < b_{n+1} \leq b_n$ for all $n \geq 1$, then $\cap_{n=1}^{\infty} I_n$ is non-empty.* □

Proof Let $a = \sup\{a_n\}$. Since $a_n \leq b_1$ for all n exists by completeness axiom, $a_n \leq b_k$ for any value of n and k, and so $a \leq b_k$. Hence $a_k \leq a \leq b_k$ for all k, and that $a \in \cap_{n=1}^{\infty} I_n$.

Let M be a set. A function $d : M \times M \to [0, \infty)$ is called a **metric** on M if

1. $d(x, y) = 0$ iff $x = y$;
2. $d(x, y) = d(y, x)$ for all $x, y \in M$;
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in M$.

The pair (M, d) is then called a **metric space**. It is easy to see any $N \subseteq M$ is also a metric space using the same d.

Example 1. On $\mathbb{R}$, $d(x, y) = |y - x|$ gives a metric.

2. On $\mathbb{R}^2$, $d_1(x, y) = |y_1 - x_1| + |y_2 - x_2|$ is also a metric, but notice that, for example, $d_1((1, 1), (0, 0)) = 2$ as opposed to the expected distance of $\sqrt{2}$.

The standard (Euclidean) metric in $\mathbb{R}^2$ is given by

We will not be distinguishing vectors by bold quantities in this document.

$$d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2}.$$

Let V be a real vector space. An **inner product** on V is a function $(\cdot, \cdot) : V \times V \to \mathbb{R}$ that, for all $x, y \in V$, satisfies the following:

- linearity in the first factor;
- $(x, y) = (y, x)$;
- $(x, x) \geq 0$ and is zero iff $x = 0$.

Example 1. For $V = \mathbb{R}^n$, the standard inner product is given by $(x, y) = x_i y_i$ (where Einstein notation is understood). If A is a symmetric matrix, then $(x, y) = x^T A y$ is an inner product if all eigenvalues of A are positive.

2. For $V = C[a, b]$, $(f, g) = \int_a^b f(x)g(x)\,\mathrm{d}x$ is an inner product since V is a vector space of continuous functions, and the only function that is everywhere zero and continuous is $f(x) = 0$ for all $x \in [a, b]$.

Theorem 1.1.2 (Cauchy–Schwartz inequality) *Let V be a real vector space, and $(\cdot, \cdot)$ an inner product on V. Then*

$$|(x, y)| \leq \|x\| \cdot \|y\|,$$

where $\|\cdot\|$ is the standard Euclidean norm of the vector, and there is equality iff $x = \lambda y$ for some $\lambda \in \mathbb{R}$.

Proof Note that $(x, 0) = (x, x - x) = (x, x) - (x, x) = 0$, so we may assume that $y \neq 0$. Then, with $\lambda = -(x, y)/\|y\|^2$,

$$\begin{aligned} 0 \leq (x + \lambda y, x + \lambda y) &= \|x\|^2 + 2\lambda(x, y) + \lambda^2 \|y\|^2 \\ &= \|x\|^2 - \frac{(x, y)^2}{\|y\|^2}. \end{aligned}$$

So $(x, y)^2 \leq \|x\|^2 \|y\|^2$ and the result follows. ∎

Lemma 1.1.3 *Let V be a real vector space with inner product $(\cdot, \cdot)$. Then $d : V \times V \to [0, \infty)$ with $d(x, y) = \|x - y\|$ gives a metric on V.*

Proof Clearly $d(x, x) = 0$ and is symmetric, so we just need to check the triangle inequality. By Cauchy–Schwartz,

$$\begin{aligned} \|a + b\| &= \sqrt{\|a\|^2 + 2(a, b) + \|b\|^2} \\ &\leq \sqrt{\|a\|^2 + 2\|a\|\|b\| + \|b\|^2} \\ &\leq \|a\| + \|b\|, \end{aligned}$$

as required. ∎

Let $f : M \to N$ be a function metric metric spaces (M, d_M) and (N, d_N). For $a \in M$, f is **continuous at** a if, for all $\epsilon > 0$, there exists $\delta > 0$ such that $d_N(f(a), f(x)) < \epsilon$ for all $x \in M$ when $d_M(a, x) < \delta$.

1.2 *Sequences and Cauchy sequences*

Let M be a metric space. A **sequence** (a_n) in M consists of elements $a_n \in M$ for all $n \in \mathbb{N}$. Let $a \in M$, and (a_n) **converges to** a if, for all $\epsilon > 0$, $d(a_n, a) < \epsilon$ for some all $n \geq n_0$. We write $\lim_{n\to\infty} a_n = a$. The sequence (a_n) is called **convergent** if there exists $a \in M$ where $a_n \to a$.

Lemma 1.2.1 *Let $f : M \to N$ be a function between metric spaces and $a \in M$. The function f is continuous at $a \in M$ iff $f(a_n) \to f(a)$ for $(a_n) \in M$ with $a_n \to a$. (Note that $f(a_n)$ is a sequence in N.)*

Proof Assume that f is continuous at $a \in M$, and let (a_n) be a sequence with $a_n \to a$. By continuity, for any $\epsilon > 0$, there exists $\delta > 0$ such that, for $d(a, y) < \delta$, $d(f(a), f(y)) < \epsilon$ for arbitrary $y \in M$. Choose $n_0 \geq 0$ such that $d(a_n, a) < \delta$ for all $n \geq n_0$, then this implies $d(f(a_n), f(a)) < \epsilon$, and thus $f(a_n) \to f(a)$ as required.

On the other hand, assume $f(a_n) \to f(a)$ for all sequences such that $a_n \to a$. Given $\epsilon > 0$, assume that instead there is no $\delta > 0$ such that, for $d(a, y) < \delta$, $d(f(a), f(y)) < \epsilon$ for arbitrary $y \in M$. Then we can find $a_n \in M$ with $d(a, a_n) < 1/n$. However, this means $d(f(a), f(a_n)) \geq \epsilon$, which contradicts the assumption that $f(a_n) \to f(a)$ even though $a_n \to a$. So such δ exists and we have continuity. ∎

Lemma 1.2.2 *The limit of a sequence is unique.*

Proof Assume there are two limits a and b for the sequence a_n. Then $d(a, b) \leq d(a, a_n) + d(a_n, b)$. As $n \to \infty$, the RHS tends to zero so $a = b$. ∎

A **Cauchy sequence** (a_n) in the metric space M is a sequence such that, for all $\epsilon > 0$, there exists $n_0 \geq 0$ such that $d(a_p, a_q) < \epsilon$ for all $p, q \geq n_0$.

Lemma 1.2.3 *A convergent sequence is a Cauchy sequence (the converse is not true).*

Proof Suppose $a_n \to a$. Then, for all $\epsilon > 0$, there is some $n_0 \geq 0$ such that $d(a_n, a) < \epsilon/2$ for $n \geq n_0$. Let $p, q \geq n_0$, then $d(a_n, a_q) \leq d(a_p, a) + d(a_q, a) < \epsilon$, so the sequence is Cauchy. ∎

A metric space M is **complete** if all Cauchy sequences in M converges.

Theorem 1.2.4 *The real line $\mathbb{R}$ is complete.*

Proof Let (a_n) be a Cauchy sequence in $\mathbb{R}$. Define the sequence of integers (n_k) where $n_0 = 1$, and n_{k+1} is the smallest integer bigger than n_k where $|a_p - a_q| < 2^{-(k+2)}$ for $p, q \geq n_{k+1}$. Define the intervals $I_k = [a_{n_k} - 2^{-k}, a_{n_k} + 2^{-k}]$ and let $x \in I_{k+1}$. Now, since $x \in I_{k+1}$, this implies that $|x - a_{n_{k+1}}| < 2^{-(k+1)}$. By definition of the integer sequence, $|a_{n_k} - a_{n_{k+1}}| < 2^{-(k+1)}$, so then, by triangle inequality,

$$|a_{n_k} - x| \leq |x - a_{n_{k+1}}| + |a_{n_{k+1}} - a_{n_k}| < 2 \cdot 2^{-(k+1)} = 2^{-k},$$

so $x \in I_k$. However, $x \in I_{k+1}$, so $I_{k+1} \subset I_k$. By Lemma 1.1.1, $\cap_{k=1}^{\infty} I_k \neq \emptyset$, so assume $a \in \cap_{k=1}^{\infty} I_k$. For $m \geq n_k$,

$$|a - a_m| \leq |a - a_{n_k}| + |a_{n_k} - a_m| \leq 2^{-k} + 2^{-(k+1)} \to 0$$

as $m \geq n_k \to \infty$. Thus $a_m \to a$ and this arbitrary Cauchy sequence converges in $\mathbb{R}$ and thus $\mathbb{R}$ is complete. ∎

Proposition 1.2.5 *For $X \neq \emptyset$, let $\mathcal{B}(X)$ be the set of functions $f : X \to \mathbb{R}$ such that f is bounded. For $f, g, \in \mathcal{B}(X)$, let $d(f,g) = \sup_{x \in X} |f(x) - g(x)$. Then $(\mathcal{B}(X), d(f,g))$ defines a complete metric space.*

Proof d is clearly a metric. For completeness, let (f_n) be a Cauchy sequence in $\mathcal{B}(X)$. For $x \in X$, $(f_n(x))$ is a Cauchy sequence of real numbers because, by definition of $d(f,g)$, $|f_q(x) - f_p(x)| \leq d(f_p - f_q)$, and since $\mathcal{R}$ is complete, the sequence $(f_n(x))$ converges.

Defining $f : X \to \mathbb{R}$ such that $f(x) = \lim_{n \to \infty} f_n(x)$, we need to show that $f \in \mathcal{B}(X)$, and that indeed $f_n(x) \to f(x)$ regardless of $x \in X$. Be definition of a Cauchy sequence, for $\epsilon > 0$, there exists $n_0 \geq 0$ such that $d(f_p, f_q) < \epsilon/2$ for $p, q \geq n_0$. Note also that, for all $x \in X$, there exists $n_1(x) \geq n_0$ such that $|f_{n_1(x)} - f| < \epsilon/2$. Then, let $x_\in X$ and $n \geq n_0$, we have

$$|f_n(x) - f(x)| \leq |f_n - f_{n_1(x)}| + |f_{n_1(x)} - f| < \epsilon.$$

Additionally note, $|f(x)| \leq |f(x) - f_{n_0(x)}| + |f_{n_0(x)}| \leq \epsilon + c_{f_{n_0}}$ since $f_{n_0(x)}$ is bounded, so $f \in \mathcal{B}(X)$. Further, $d(f_n - f) = \sup |f_n - f| = \delta < \epsilon$, so f_n converges to $f \in \mathcal{B}(x)$. Thus every Cauchy sequence converges and thus the space is complete and equipped with a metric. ∎

1.3 *Topology of metric spaces*

Let (M, d) be a metric space with $x \in M$ and $r > 0$. Define the **open ball** around x of radius r to be

$$B(x;r) = \{y \in M \ : \ d(x,y) < r\}.$$

The analogous **closed ball** $D(x;r)$ is defined with the less than or equal to sign. A set $A \subset M$ is **bounded** if it can be contained in some

$D(x;r)$ for some $x \in M$, $r > 0$. A set $U \subset M$ is **open** if, for all $x \in U$, there exists $r_x > 0$ such that $B(x;r_x) \subset U$. A set $A \subset M$ is **closed** if $M \setminus A$ is open.

Lemma 1.3.1 *Let (M,d) be a metric space, then:*

1. *M and $\emptyset$ are open;*
2. *$\bigcup_i A_i$ is open if all $A_i \subset M$ are open;*
3. *$\bigcap_i^n$ is open if all $A_i \subset M$ are open and $n < \infty$;*
4. *$B(x;r)$ is open for some $r > 0$.*

Proof The first two are obvious. For 3), suppose the open sets U_i indexed by i are open and $x \in \bigcap_{i=1}^n U_i$. Then $xinU_i$ for all i, so there is some $B(x;r_i) \subset U_i$. Taking the minimum of such $r_i > 0$ means $B(x;r_i) \subset \bigcap_{i=1}^n U_i$, and thus the collective finite union is open.

For 4), let $y \in B(x;r)$, $r_y = r - d(x,y) > 0$ and $z \in B(y;r_y)$. Then $d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + r - d(x,y) = r$, so $B(y;r_y) \subseteq B(x;r)$. ■

Corollary 1.3.2 *The following may be shown by considering the appropriate complements:*

1. *M and $\emptyset$ are closed;*
2. *$\bigcap_i A_i$ is closed if $A_i \subset M$ for all i;*
3. *$\bigcup_i A_i$ is closed if $A_i \subset M$ for all i and $n < \infty$;*
4. *$D(x;r)$ is closed.*

□

Example Open intervals are open and closed intervals are closed.
(a,∞) is open as it is a union of open bounded intervals.
$[a,\infty)$ is closed since $(-\infty,a)$ is open.
$\mathbb{Z}$ is closed as $\mathbb{R} \setminus (\cup_{n=-\infty}^{\infty}(n,n+1))$ is closed.
$\mathbb{Q}$ and $[0,1)$ are neither, while $\mathbb{R}$ is both.

Proposition 1.3.3 *Suppose M is a metric space and $A \subseteq M$. A is closed iff every sequence converges to $a \in A$.*

Proof Assume A is closed and $a_n \to a$. Assume the converse so that $a \in U = M \setminus A$ which is an open set. Then there is some $r > 0$ such that $B(a;r) \in U$, and since $a_n \to a$, there exists $n_0 \geq 0$ where $d(a_n,a) < r$ for $n \geq n_0$. This implies $a_n \in B(a;r)$ for all n, but this is a contradiction since $a_n \in A$, and thus $a \in A$.

Assume $a_n \to a \in A$. Let $x \in M \setminus A$, $r > 0$, and assume there is no such $B(x;r) \subset M \setminus A$. Thus there is an intersection, i.e., $B(x;1/n) \cap A \neq \emptyset$. This implies that there is some i where $a_i \in B(x;1/n) \cap A$. However, (a_n) is a sequence in A and $d(a_m, x) < 1/n$ for $m \geq n+1$, so $a_m \to a$, but this implies $x = a$ which is not possible since $x \in M \setminus A$. So $M \setminus A$ is open which means A is closed. ■

Theorem 1.3.4 *Let M be a complete metric space and $A \subseteq M$ is closed. Then A is complete with the induced metric.*

Proof Let (a_n) be a Cauchy sequence in A. Since M is complete, (a_n) converges in M, but A is closed, so (a_n) converges in A by previous proposition, which implies A is complete. ■

Let M be a metric space. M is **compact** if every sequence $(a_n) \in M$ has a convergent subsequence (a_{n_k}).

Example • $(a_n) = (-1)^n$ is non-convergent but has a convergent sequence.

- $M = (0,1)$ is not compact since $a_n = 1/n$ and its subsequences do not converge in M.
- $\mathbb{R}$ is not compact as a_n has no subsequence converging in $\mathbb{R}$.
- $M = [0,1]$ is compact. Let (a_n) be a subsequence in M. Let I_1 be either $[0, 1/2]$ or $[1/2, 1]$, and let (a_{n_k}) be the subsequences in I_1. Continuing this we have a sequence of intervals $I_{m+1} \subset I_m$ with I_m of length 2^{-m}. Denote the subsequences $(a^m_{m_k})$ to be those in I_m. Taking $b_m = a^m_{n_m} \in I_m$, we see that $b_{m+1} \in I_M$ since $I_{m+1} \subset I_m$, so that $d(b_m, b_q) \leq 2^{-m}$ for $q \geq m$. Thus (b_m) is a Cauchy sequence, which is a subsequence of (a_n). Since $M \subseteq \mathbb{R}$, M is complete, so $b_m \to b \in M$, and thus M is compact.

Proposition 1.3.5 *By extension, closed n-gons in $\mathbb{R}^n$ are compact.* □

Proposition 1.3.6 *Let $f : M \to N$ be a continuous map between metric spaces. If M is compact, then $f(M) \subset N$ is compact.*

Proof Let (a_n) be a sequence in $f(M)$. Then $a_n = f(b_n)$ for some $b_n \in M$. The sequence (b_{n_k}) converges in M since M is compact, thus

$$\lim_{k\to\infty} a_{n_k} = \lim_{k\to\infty} f(b_{n_k}) = f\left(\lim_{k\to\infty} b_{n_k}\right) = f(b)$$

since f is continuous. So (a_{n_k}) is convergent, thus $f(M)$ is compact. □

Proposition 1.3.7 *A closed subset of a compact space is a compact set.*

Proof Let (a_n) be a sequence in $A \subset M$ where M is compact. Since $(a_n) \in M$, (a_{n_k}) is convergent, but A closed so $(a_{n_k}) \to a \in A$, thus A is compact. □

1.3.1 Heine–Borel theorem

Theorem 1.3.8 *A subset $A \subseteq \mathbb{R}^n$ is compact iff A is closed and bounded.*

Proof Suppose A is compact, so clearly A is closed. If A is unbounded, then there exists $(a_n) \in A$ where $d(a_n, 0) \geq n$, so (a_{n_k}) does not converge in $\mathbb{R}^n$. However A is compact, which is a contradiction, so A is bounded.

Suppose A is bounded, then $A \subseteq [a, b]^n$. If A is closed, then it is a closed subset of a compact set, so A is compact by previous proposition. ■

For example, if $f : M \to N$ with f is a scalar continuous function, then $f(M) \subset \mathbb{R}$ is closed and bounded since M is compact, and thus $f(M)$ compact implies $f(M)$ is closed and bounded.

1.4 Banach and Hilbert spaces

Let V be a real vector space. The **norm** on V is a function $\|\cdot\| : V \to [0, \infty)$ where:

1. $\|x\| = 0$ iff $x = 0$;
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$ for all $x \in V$ and $\lambda \in \mathbb{R}$;
3. $\|x + y\| \leq \|x\| + \|y\|$.

The pair $(V, \|\cdot\|)$ gives a **normed vector space**.

Lemma 1.4.1 *Let V be a normed vector space, then $d(x, y) = \|x - y\|$ defines a metric on V.*

Proof Two of the properties follow from definition. To show the reflexive property, note that

$$d(y, x) = \|y - x\| = \|(-1)(x - y)\| = \|x - y\| = d(x, y).$$

■

Example 1. It may be shown that the metrics

$$\sum_i |x_i|, \qquad \sum_i \sqrt{|x_i|^2}, \qquad \max\{|x_i| \in \mathbb{R}\}$$

define norms on $\mathbb{R}^n$ (the ℓ^1, ℓ^2 and ℓ^∞ norms).

2. The **supremum norm** on $B(X)$ is defined by

$$\|f\|_\infty = \sup\{|f(x)| \in \mathbb{R} \,;\, x \in X\}.$$

3. For X a metric space, $C_b(X) = \{f : x \to \mathbb{R} \ : \ f$ continuous and bounded$\}$ is also a normed vector space with the supremum norm.

If $C(X) = \{f : x \to \mathbb{R} \ : \ f \text{ continuous}\}$ then f does not have a supremum, however, we have the following:

Proposition 1.4.2 *If X is compact, then $C(X) = C_b(X)$, so $C(X)$ is a normed vector space.*

Proof $C_b(X) \subseteq C(X)$ regardless of X. For the converse, assume $f \in C(X)$, so that $f(X)$ is compact. This implies $f(X)$ is bounded and closed by the Heine–Borel theorem, so $C(X) \subseteq C_b(X)$. ∎

Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. A function $f : V \to W$ is continuous at $x \in V$ if, for all $\epsilon > 0$, there exists $\delta > 0$ such that $\|x - y\|_V < \delta$ implies that $\|f(x) - f(y)\|_W < \epsilon$.

Let V be a normed vector space. V is a **Banach space** if V with the metric induced by the norm is complete.

Theorem 1.4.3 *Let X be a metric space, then $C_b(X)$ with the supremum norm is a Banach space.*

Proof Since $C_b(X) \subseteq B(X)$, if C_b is closed, then C_b is complete since $B(X)$ is complete. To show this, let $(f_n) \in C_b(X)$, and let $f_n \to f \in B(X)$. The convergene of f_n implies that there exists $n_0 \geq 0$ such that $\|f_n - f\| < \epsilon/3$ for any $\epsilon > 0$ with $n \geq n_0$. Also, $\|f_{n_0}(y) - f(y)\| < \epsilon/3$ for all $y \in X$. The functions are continuous, so there exists $\delta > 0$ where, if $d(x, y) < \delta$, $\|f_{n_0}(x) - f_{n_0}(y)\| < \epsilon/3$ for $x \in X$. Thus, for $d(x, y) < \delta$,

$$|f(x) - f(y)| \leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(y)| + |f_{n_0}(y) - f(y)| < \epsilon,$$

so f is continuous, and $C_b(X)$ is closed and thus complete. ∎

Corollary 1.4.4 *For $a < b$, $C[a, b]$ with the supremum norm is a Banach space.* □

Note that $C[a, b]$ is not a complete space with, for example, the L_2 norm

$$\|f\|_2 = \sqrt{\int_a^b (f(x))^2 \, dx}.$$

For example, with $f_n = x^n$, $f_n \to 0$ but clearly $f_n(1) = 1$ for all n. The underlying reason is the sequence is not a Cauchy sequence with respect to the norm.

Convergence with respect to the supremum norm is called **uniform convergence** (cf. Complex Analysis 2H).

Let $(V, \|\cdot\|)$ be a Banach space. If there is an inner product from V which induces this norm, then V is called a **Hilbert space**.

Theorem 1.4.5 *Let (M, d) be a metric space. Then there exists $(\overline{M}, \overline{d})$ where $\overline{M}$ is complete, and there is an embedding $\iota : M \to \overline{M}$ with $d(x, y) = d(\iota(x), \iota(y))$ for all $x, y \in M$. Also, for all $\overline{x} \in \overline{M}$, there is a sequence $(x_n) \in M$ with $x_n \to \overline{x}$ as $n \to \infty$.* $\square$

Here, $\overline{M}$ is called the **completion** of M, and it is unique up to some isomorphism.

Example The completion of $\mathbb{Q}$ is $\mathbb{R}$ with respect to the Euclidean metric.

The completeness of $C[a, b]$ with respect to the inner product metric is denoted $L^2[a, b]$..

Note that elements of L^2 are not exactly functions, but rather *equivalence classes* (cf. $11 \equiv 1$ modulo 10)

1.4.1 *The contraction mapping theorem*

Theorem 1.4.6 *Let (M, d) be a complete metric space, $0 \leq \lambda \leq 1$ and a $f : M \to M$ with $d(f(x), f(y)) \leq \lambda d(x, y)$ for all $x, y \in M$. Then f has one unique fixed point where $f(x_0) = x_0$.*

Or, if you throw a map of the world on the floor, there is exactly one point on the map that exactly corresponds to one point on the floor.

Proof Note that f is a contraction, and continuity is automatically satisfied from the condition that $d(f(x), f(y)) \leq \lambda d(x, y)$.

Let $x \in M$, and $a_n = f^n(x)$. So we have

$$\begin{aligned} d(x, a_n) &\leq d(x, f(x)) + d(f(x), f^2(x)) + \ldots d(f^{n-1}(x), f^n(x)) \\ &= \sum_{i=0}^{n-1} d(f^i(x), f^{i+1}(x)) \\ &\leq \sum_{i=0}^{n-1} \lambda d(x, f(x)) \\ &= d(x, f(x)) \frac{1 - \lambda^n}{1 - \lambda} \\ &\leq \frac{d(x, f(x))}{1 - \lambda}, \end{aligned}$$

by Cauchy–Schwartz and the arithmetic progression with $0 \leq \lambda < 1$. Now,

$$d(a_n, a_m) = d(f^n(x), f^m(x)) \leq \lambda^m d(f^{n-m}, x) \leq \lambda^m \frac{d(x, f(x))}{1 - \lambda}$$

assuming $n > m$. For $n, m \geq n_0$, we have

$$d(a_n, a_m) \leq \lambda^{n_0} \frac{d(x, f(x))}{1 - \lambda}.$$

Clearly (a_n) is a Cauchy sequence, and thus we have completeness and $a_n \to a \in M$. Now,

$$f(a) = f\left(\lim_{n\to\infty} a_n\right) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} a_{n+1} = a,$$

so there is some $a \in M$ that is a fixed point.

To show uniqueness, suppose b is another fixed point. Then

$$d(a,b) = d(f(a), f(b)) \leq \lambda d(a,b),$$

and for $\lambda \neq 0$, $d(a,b) = 0$, so $a = b$. ■

1.5 *A norm for matrix spaces*

We want a norm reflecting the fact that matrices can be identified with linear maps. Let $A = (A_{ij}) \in M_{n,k}(\mathbb{R})$. We define

$$\|A\| = \sup\{\|Ax\|_2 \ : \ x \in \mathbb{R}^k, \ \|x\|_2 \leq 1\}, \tag{1.1}$$

where $\|\cdot\|$ is the Euclidean norm. Here, $Ax \in \mathbb{R}^n$, and $x \mapsto \|Ax\|_2$ is clearly a continuous map. By the Heine–Borel theorem, $\{\|Ax\|_2 \ : \ \|x\|_2 \leq 1\}$ is bounded and closed, so the supremum exists, and there is x with $\|x\|_2 \leq 1$ such that $\|A\| = \|Ax\|_2$ exists.

Lemma 1.5.1 *We have*

- $\|Ax\|_2 \leq \|A\|\|x\|_2$ *for all* A *and* x
- $\|AB\| \leq \|A\|\|B\|$
- $\|A\|_\infty \leq \|A\| \leq k\sqrt{n}\|A\|_\infty$,

where $\|A\|_\infty = \max\{|A_{ij}| \ : \ A \in M_{n,k}(\mathbb{R})\}$. □

2 Ordinary differential equations

Let $U \subset \mathbb{R}^n$ be open. A **vector field** or **autonomous differential equation** is a continuous map $v : U \to \mathbb{R}^n$ with no explicit time dependence. Here, U is called the **phase space** of v.

For $x \in U$, $\tau \in \mathbb{R}$, a continuous differential curve $\alpha : (a, b) \to U$ is an **integral curve** of v at (x, τ) if $\tau \in (a, b)$, $\alpha(t) = x$ and $\alpha'(t) = v(\alpha(t))$. Note the integral curves have tangent vectors which agree with v at a given point.

More generally, for $U \in \mathbb{R}^n$, $I \subset \mathbb{R}$, a **differential equation** is a continuous map $V : U \times I \to \mathbb{R}^n$. A **solution** of V at $x \in U$ and $\tau \in I$ is a continuously differential curve $\alpha : I \to U$ with $\alpha'(t) = V(\alpha(t), t)$ and $\alpha(t) = x$.

2.1 Picard–Lindelöf theorem

This is an existence and uniqueness theorem for differential equations.

Theorem 2.1.1 *Let $U \subset \mathbb{R}^n$, $I \subset \mathbb{R}$ be open and $V : U \times I \to \mathbb{R}^n$ be a differential equation where, for all $x_1, x_2 \in U$, $t \in I$, there exists $L \geq 0$ such that*

$$\|v(x_1, t) - v(x_2, t)\| \leq L\|x_1 - x_2\|_2.$$

Given $(u, \tau) \in U \times I$, there exists $a, b > 0$ with

$$U_1 = \{x \in U \ : \ \|x - u\| < a\}, \quad I_1 = \{t \in I \ : \ |t - \tau| < b\}$$

such that the differential equation v has an unique solution for all $x \in U_1$ and $\tau \in I_1$. Furthermore, the resulting $\alpha : U_1 \times I_1 \to U$ given by $\alpha(x, t) = \alpha_x(t)$ is continuous.

Compare this with the **Lipschitz condition** where $\|v(x_1) - v(x_2)\| \leq L\|x_1 - x_2\|$, where L is the **Lipschitz constant**.

Proof This one is quite long! The key idea is to construct a contraction mapping A and make use of the fixed point theorem to demonstrate existence and uniqueness. We are going to split this up into little bits.

- We first construct an integral curve α with $\partial\alpha/\partial t(x, t) = v(\alpha, t)$,

$\alpha(x,\tau) = x$. By integrating,

$$\alpha(x,t) = x + \int_\tau^t v(\alpha(x,s),s)\ \mathrm{d}s.$$

Define some operator A such that

$$A\beta(x,t) = x + \int_\tau^t v(\beta(x,s),s)\ \mathrm{d}s,$$

then we note that $A\alpha = \alpha$, and α is a fixed point of the operator A. We aim to show that A is a contraction in a space satisfying the relevant properties.

- Let $a_1, b_1 > 0$ be such that

  $$D_1 = D(u; 2a_1) \subset U, \quad D_2 = D(\tau; b_1) \subset I.$$

  By the Heine–Borel theorem, $D_1 \times D_2 \subset \mathbb{R}^{n+1}$ is compact, and so there exists some $K \geq 0$ such that, with respect to the Euclidean norm, $\|v(x,t)\| < K$ for all $(x,t) \in D_1 \times D_2$.

  Recall D denotes *closed* balls, while B denote open balls.

  Let $a, b > 0$ be such that

  $$0 < a < a_1, \qquad b < \min\left\{b_1, \frac{a}{K}, \frac{1}{L}\right\}.$$

  Recall that $U_1 = B(u; a)$ and $I_1 = B(\tau; b)$, so let

  $$M = \{\beta \ : \ U_1 \times I_1 \to D \subset \mathbb{R}^n\}$$

  where β is continuous and $\beta(x,\tau) = x$ for all $x \in U_1$. This implies that

  $$M \subseteq \left(C_b(U_1 \times I_1)\right)^n,$$

  and since $\left(C_b(U_1 \times I_1)\right)^n$ is a Banach space with the supremum norm, if M is closed, then M is complete.

- Suppose $(\beta_n) \in M$ where $\beta_n \to \beta$. For $(x,t) \in U_1 \times I_1$, $\|\beta(x,t) - \beta_n(x,t)\| \leq \|\beta - \beta_n\|$ so $\beta_n \to \beta$, but since D_1 is closed, $\beta \in D$ and obviously $\beta_n((x,\tau) \to \beta(x,\tau) = x$, so M is closed and so is complete.

- If we now consider $A\beta$, then we have $A\beta(x,\tau) = x$ and that

  $$\begin{aligned}
  \|A\beta(x,t) - u\| &\leq \|A((x,t) - x\| - \|x - u\| \\
  &\leq \int_\tau^t \|v(\beta(x,s),s\|\ \mathrm{d}s + a \\
  &\leq K|t - \tau| + a \\
  &\leq Kb + a \\
  &< 2a < 2a_1
  \end{aligned}$$

  by Cauchy–Schwartz, definition of U_1, second bullet point, definition of I_1, and definition of b and a respectively. By definition of D_1, we have $A\beta(x,t) \in D_1$.

- Note then we have

$$\begin{aligned}
\|A\beta(x,t) - A\beta(y,t') &\leq \|x-y\| + \left\| \int_\tau^t v(\beta(x,s),s) - v(\beta(y,s),s) \, \mathrm{d}s \right\| \\
&\quad + \left\| \int_t^{t'} v(\beta(y,s),s) \, \mathrm{d}s \right\| \\
&\leq \|x-y\| + L \int_\tau^t \|v(\beta(x,s),s) - v(\beta(y,s),s)\| \, \mathrm{d}s \\
&\quad + K|t-t'| \\
&\leq \|x-y\| + L \sup_{s\in[\tau,t]} \|\beta(x,s) - \beta(y,s)\| + K|t-t'|,
\end{aligned}$$

  by the Lipschitz conditions. All terms can be made arbitrarily small since x can be made close to y, t can be made close to t', and since $[\tau, t]$ is compact, $\|\beta(x,s) - \beta(y,s)\|$ can be made arbitrarily small. So now $A\beta \in D_1$ is continuous, and therefore $A\beta \in M$, and $A : M \to M$ is a self mapping.

- Since A is a self-mapping, for $\beta_{1,2} \in M$, we have

$$\begin{aligned}
\|A\beta_1 - A\beta_2\| &\leq \int_\tau^t \|v(\beta_1(x,s),s) - v(\beta_2(x,s),s)\| \, \mathrm{d}s \\
&\leq L \int_\tau^t \|\beta_1 - \beta_2\| \, \mathrm{d}s \\
&= L|t-\tau|\|\beta_1 - \beta_2\| \\
&\leq (Lb)\|\beta_1 - \beta_2\|
\end{aligned}$$

  by definition of I_1. Note that $Lb < 1$ by the definition of b, and therefore A is a contraction.

  $b < 1/L$.

Since A is a contraction and M is complete, by contraction mapping there is one unique point in M that is fixed under A. Clearly this is α by definition of β (see first bullet point), and hence α is the unique solution to the ODE satisfying the stated conditions. ■

Note that it doesn't matter if $\alpha : I_1 \to U$, since we can redefine M and A as $M_x = \{\beta : I_1 \to D\}$ with $\beta(t) = x$, and $A_x : M_x \to M_x$. There will be an unique solution for fixed $x \in U_1$, where the generation solution gives this solution.

2.2 *Differentiation in $\mathbb{R}^n$*

Let $U \subset \mathbb{R}^n$ be open. Recall that $f : U \to \mathbb{R}^n$ is differentiable at $x \in U$ with derivative

$$Df(x) = \left(\frac{\partial f_i}{\partial x_j} \right) \in M_{p,n}(\mathbb{R}) \tag{2.1}$$

if near x we can write

$$f(x+h) = f(x) + Df(x) \cdot h + R(h), \qquad \lim_{\|h\| \to 0} \frac{R(h)}{\|h\|} = 0.$$

If f is differential for all $x \in U$, then $Df : U \to M_{p,n}(\mathbb{R}) = \mathbb{R}^{pn}$. If $D^i f$ is continuous then f is said to be of i-**class**, with $f \in C^i(U)$.

2.2.1 Mean value theorem

Theorem 2.2.1 *Let $U \subset \mathbb{R}^n$ be open, $x \in U$, $h \in \mathbb{R}^n$ where $x + th \in U$ for all $t \in [0,1]$ and $f \in C^1 : U \to \mathbb{R}^p$, then*

$$f(x+h) - f(x) = \int_0^1 Df(x+th) \cdot h \, \mathrm{d}t.$$

Proof Let $f_i : U \to \mathbb{R}$ with $g_i(t) = f_i(x+th)$, so that $g : [0,1] \to \mathbb{R}$. Then we have $g_i'(t) = Df_i(x+th) \cdot h$. By the fundamental theorem of calculus,

$$\begin{aligned} g_i(1) - g_i(0) &= \int_0^1 Df_i(x+th) \cdot h \, \mathrm{d}t \\ &= f_i(x+th) - f_i(x). \end{aligned}$$

Since this is true per component, we have the result in higher dimensions. ■

Corollary 2.2.2 *Let $U \subset \mathbb{R}^n$ be open and convex*[1]*, and also that $f \in C^1 : U \to \mathbb{R}^n$. Assume that there exists some $C = \sup\{\|Df(x)\| \in \mathbb{R} : x \in U\}$, then $\|f(y) - f(x)\| \leq C\|y - x\|$.*

[1] So for all $x, y \in U$, $xt + (1-t)y \in U$ for $t \in [0,1]$.

Proof By the mean value theorem, we have

$$\begin{aligned} \|f(x+h) - f(x)\| &\leq \int_0^1 \|Df(x + h \cdot h\| \, \mathrm{d}t \\ &\leq \int_0^1 \|Df(x+h)\| \cdot \|h\| \, \mathrm{d}t \\ &\leq \int_0^1 C \cdot \|h\| \, \mathrm{d}t = C \cdot \|h\|. \end{aligned}$$

Since h is arbitrary (up to us assuming convexity), letting $h = y - x$ leads the result. ■

Note that for the above corollary, U can always be reduced so that C exists locally. For the Picard–Lindelöf theorem, we get $v \in C^1 : U \times I \to \mathbb{R}$ implies the Lipschitz condition is satisfied locally.

2.2.2 Matrices

Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^p$ be open. A C^1-function $f : U \to V$ is a **diffeomorphism** if there exists $f^{-1} : V \to U$ where $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}$ (the identity map), and f^{-1} is differential for all $x \in V$.

Example $f = x^3$ has $f^{-1} = x^{1/3}$, but since f^{-1} is not differentiable at $x = 0$, x^3 is not a diffeomorphism on $\mathbb{R}$.

By the chain rule, note that

$$D(f^{-1} \circ f) = \left(Df^{-1}(f)\right) Df = I_n, \quad D(f \circ f^{-1}) = \left(Df(f^{-1})\right) Df^{-1} = I_p.$$

If $y = f(x)$ then $Df^{-1}(y) = (Df(x))^{-1}$, then inverse matrix of $Df(x)$, so $Df(x)$ is invertible and $p = n$ if f is a diffeomorphism.

Lemma 2.2.3 *1. $GL_n(\mathbb{R})$ is an open set.*

This is the general linear group with real entries.

2. *$A \in M_{n,n}(\mathbb{R})$ with $\|A\| \leq 1$ implies that $I - A \in GL_n(\mathbb{R})$.*

3. *$inv : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ with $A \mapsto A^{-1}$ is a smooth diffeomorphism.*

Proof Recall that the determinant is defined as

S_n here is the group of symmetric permutations, and $\mathrm{sig}(\sigma)$ is the signature of the permutation σ ($+1$ if even and -1 if odd).

$$\det A = |A| = \sum_{\mathrm{sig} \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

This is a polynomial in components of A, so it is a smooth function.

1. $\mathrm{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R} - \{0\})$ so $A \in \mathrm{GL}_n(\mathbb{R})$ implies that $|A| \neq 0$, which implies $|B| \neq 0$ for B close to A, and thus $\mathrm{GL}_n(\mathbb{R})$ is open.

2. If $\|A\| \leq 1$, define $B_n = \prod_{i=0}^{n} A^i$ where $A^0 = I$. $\{B_n\}$ is a Cauchy sequence since

$$\|B_n - B_m\| \leq \sum_{k=\min\{m,n\}+1}^{\max\{m,n\}} \|A\|^k \leq \frac{\|A\|}{1 - \|A\|} \to 0$$

for sufficiently large m, n with $\|A\| \leq 1$. So there exists $B = \lim_{n\to\infty} B_n$, and thus

$$(I - A)B = (I - A) \lim_{n\to\infty} B_n.$$

B_n continuous implies that

$$(I - A)B = \lim_{n\to\infty} (I - A)B_n = \lim_{n\to\infty} I - A^{n+1} = I$$

since $\|A\| \leq 1$, so $B^{-1} = I - A \in \mathrm{GL}_n(\mathbb{R})$.

3. By Cramer's rule, for $A = (a_{ij})$, $A^{-1} = (b_{ij})$ with $b_{ij} = \det A_{ij} / \det A$, where A_{ij} is the matrix obtained by replacing the i^{th} column with the standard j^{th} basis vector. So (b_{ij}) depends smoothly on (a_{ij}) since det is a smooth map, and so inv is smooth. Note additionally that $\mathrm{inv} \circ \mathrm{inv} = \mathrm{id}$, so it is a bijection and hence a diffeomorphism.

■

2.2.3 Inverse function theorem

Let $U \subset \mathbb{R}^n$ be open and $f \in C^k : U \to \mathbb{R}^n$. f is **locally invertible** at $x \in U$ if there exists $U_1 \subset U$ such that for $x \in U_1$, $V_1 \subset \mathbb{R}^n$ where $f(x) \in V_1$ is open and $f : U_1 \to V_1$ is a diffeomorphism.

Theorem 2.2.4 *Let $U \subset \mathbb{R}^n$ be open and $f \in C^k : U \to \mathbb{R}^n$, $u \in U$. f is locally invertible iff $Df(u)$ is invertible. Here the local inverse is of class C^k.*

Proof This one is quite long!

- If f is locally invertible at u, then it is a diffeomorphism, so clearly $Df(u)$ is invertible. However, this is for an isolated point, and we need to show that is also true on the appropriate neighbourhood.

- Assume that $u = 0 = f(u)$, i.e. a fixed point, and $Df(0) = I$. Define, for $y \in \mathbb{R}^n$,

  $$g_y(x) = y + x - f(x) \quad \Rightarrow \quad y - f(x) = g_y - x.$$

  Note that $Dg_y(x) = I - Df(x)$ and does not depend on y. Also that $Dg_y(0) = I - I = 0$.

  By continuity, we have $\|Dg_y(x) = \|Dg_0(x)\| \le 1/2$ for some x near 0. This implies that

  $$\|g_y(x_1 - g_y(x_2)\| \le \frac{1}{2}\|x_1 - x_2\|$$

  for $x_{1,2} \in D(0;r)$. Taking $x_2 = 0$, we also get

  $$\|g_y(x) - y\| \le \frac{1}{2}\|x\|,$$

  so we have

  $$\|g_y(x)\| \le \frac{1}{2}\|x\| + \|y\|$$

  for $y \in D(0;r/2)$ and $x \in D(0;r)$, and thus $\|g_y(x)\| \le r$. Hence we have $g_y(x) : D(0;r) \to D(0;r)$, and $g_y(x)$ is by construction a contraction since $\|g_y(x_1 - g_y(x_2)\} \le (1/2)\|x_1 - x_2\|$.

- By contraction mapping theorem, for all $y \in D(0;r/2)$, there exists a unique $x \in D(0;r)$ with $y = f(x)$, so there exists an inverse function defined on $D(0;r/2)$.

  Define

  $$U_1 = \{x \in U \;:\; \|x\| < r, \|f(x)\| < r/2\}, \quad V_1 = f(U_1) = B(0;r/2).$$

  By definition, both the domain and image are open sets. $f : U_1 \to V_1$ is a restricted bijection since it is a bijection on $D(0;r/2) \supset$

$B(0;r/2)$. Given $x_{1,2} \in D(0;r)$, we have

$$\begin{aligned}\|x_1 - x_2\| &= \|g_0(x_1) + f(x_1) - g_0(x_2) + f(x_2)\| \\ &\leq \|g_0 - g_0(x_2)\| + \|f(x_1) - f(x_2)\| \\ &\leq \frac{1}{2}\|x_1 - x_2\| + \|f(x_1) - f(x_2)\|,\end{aligned}$$

so that $\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|$. For $x_2 = 0$, we have $\|x_1\| \leq 2\|f(x_1)\|$. Since $\|f(x_1)\| < r/2$ by construction, we have $\|x_1\| < r$, so indeed $V_1 = B(0;r/2)$.

For $f^{-1} = \phi$, $\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|$ implies that $\|\phi(y_1) - \phi(y_2)\| \leq 2\|x_1 - x_2\|$, so that f^{-1} is Lipschitz continuous.

- Note that $Df(x)$ is invertible for all $x \in D(0;r)$, since we have $g_0(x) - x = f(x)$, so that $Df(x) = I - Dg_0(x)$, but $\|Dg_0(x)\| \leq 1/2$ from point 2 above, so $Df(x)$ is invertible for all $x \in D(0;r)$, and in particular for $x \in B(0;r) \subset D(0;r)$.

- Recall that if f id differentiable, then $f(x_1) - f(x_2) = Df(x_1)(x_1 - x_2) + R(x_1 - x_2)$ with $R(h)/\|h\| \to 0$ as $\|h\| \to 0$. Let $y_i = f(x_i)$. For $i = 1, 2$,

$$y_1 - y_2 = Df(x_1)\left(\phi(y_1) - \phi(y_2)\right) + R\left(\phi(y_1) - \phi(y_2)\right),$$

so that

$$\begin{aligned}\left(Df(\phi(y_1))\right)^{-1}(y_1 - y_2) &= \left(\phi(y_1) - \phi(y_2)\right) \\ &\quad + \left(Df(\phi(y_1))\right)^{-1} R\left(\phi(y_1) - \phi(y_2)\right).\end{aligned}$$

We want to show that the remainder term tends to zero, which will show that $\phi = f^{-1}$ is differentiable. For that, note we have, by Cauchy–Schwartz and point 3 above,

$$\frac{\|\left(Df(\phi(y_1))\right)^{-1} R\left(\phi(y_1) - \phi(y_2)\right)\|}{\|y_1 - y_2\|} \leq \frac{\|\left(Df(\phi(y_1))\right)^{-1}\| \cdot \|R\left(\phi(y_1) - \phi(y_2)\right)\|}{(1/2)\|\left(\phi(y_1) - \phi(y_2)\right)\|}.$$

$\left(Df(\phi(y_1))\right)^{-1}$ is bounded since f is differentiable. Further more, f differentiable means $\|R\left(\phi(y_1) - \phi(y_2)\right)\| / \|\left(\phi(y_1) - \phi(y_2)\right)\| \to 0$ as $\|\left(\phi(y_1) - \phi(y_2)\right)\| \to 0$. Thus the desired remainder goes to zero since $y_1 - y_2 \to 0$ implies $\phi(y_1) - \phi(y_2) \to 0$, and $\phi = f^{-1}$ is differentiable.

- The derivative $D\phi(y) = \left(Df(\phi(y))\right)^{-1} = \text{inv} \circ Df \circ \phi)y$, so by construction, $D\phi = Df^{-1}$ is continuous. By chain rule, if $f \in C^k$, $D^{k-1}\phi$ is continuous, and thus $\phi = f^{-1} \in C^k$.

■

2.2.4 Implicit function theorem

Theorem 2.2.5 *Let $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ be open and $f : U \times V \to \mathbb{R}^m$ be a C^k-function, with $k \geq 1$. Let $(u, v) \in U \times V$ such that the matrix $[\partial f_i / \partial x_j](u, v)$ is invertible with $\boldsymbol{c} = f(u, v)$. Then there is a C^k-function $\eta : U_1 \to V_1$ with $u \in U_1 \subset U$, $v \in V_1 \subset V$ where $\eta(u) = v$ and $f(x, \eta(x)) = \boldsymbol{c}$ for all $x \in N((u, v); r)$. Further more, if $f(x, y) = \boldsymbol{c}$ for $(x, y) \in U_1 \times V_1$, then we have $y = \eta(x)$ in the respective sets.*

Proof Define $\phi : U \times V \to \mathbb{R}^n \times \mathbb{R}^m$ where $(x, y) \mapsto (x, f(x, y))$. We have

$$D\phi(u, v) = \begin{pmatrix} I & 0 \\ \partial f_i / \partial x_j(u, v) & \partial f_i / \partial x_j(u, v) \end{pmatrix},$$

so $\det D\phi(u, v) \neq 0$, and so by the inverse function theorem, ϕ is locally a diffeomorphism.

Since $\phi(x, y) = (x, f(x, y)$, we have $\phi^{-1}(A, \boldsymbol{b}) = (A, g(A, \boldsymbol{b})$. Setting $\eta(x) = g(x, \boldsymbol{c})$, then defining $\hat{p}_2$ as the projection operator for the second argument, we have

$$\begin{aligned} f(x, \eta(x)) &= f(x, g(x, c)) \\ &= \hat{p}_2 \phi(x, g(x, c) \\ &= \hat{p}_2 \phi \phi^{-1}(x, \boldsymbol{c}) \\ &= \hat{p}_2(x, \boldsymbol{c}) = \boldsymbol{c}. \end{aligned}$$

So we have $f(x, y) = \boldsymbol{c}$ iff $y = \eta(x)$ for $(x, y) \in W$ where ϕ is a diffeomorphism. This is achieved by choosing $u \in U_1 \subset U$, $v \in V_1 \subset V$ so that $U_1 \times V_1 \subset W$, with $\eta(U_1) = V_1$. ■

The implicit function theorem gives a criterion of when we can solve $f(x, y) = \boldsymbol{c}$ unique for y. In fact, if the linear equation $[Df(u, v)](x, y) = 0$ is uniquely solvable, then $f(x, y) = \boldsymbol{c}$ is uniquely solvable for y.

2.2.5 Manifolds

Let $M \subset \mathbb{R}^n$, $k \geq 0$, $\ell \geq 1$. M is a C^ℓ **k-dimensional manifold** if, for all $p \in M$, we also have $p \in U \subset \mathbb{R}^n$ where there exists a C^ℓ-diffeomorphism $h : U \to U' \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ with $h(U \cap M) = U' \cap (\mathbb{R}^k \times \{0\})$. Informally, a manifold is a structure where every point of M has a neighbourhood that resembles $\mathbb{R}^k$. h here is called a **chart**, which maps neighbourhoods of the manifold to $\mathbb{R}^k$ (think **co-ordinate system** or segments of maps). A collection of charts that spans the whole of M is called an **atlas**.

Notice then in the previous proof, ϕ is a chart, and $W \cap \{(x, y) \in \mathbb{R}^n : f(x, y) = c\}$ is a k-manifold.

Example An open subset $U \subset \mathbb{R}$ is a C^∞ n-manifold where the chart is $\mathrm{id} : U \to U$.

For a slightly less trivial example, consider the **unit n-sphere** $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$. With $f : \mathbb{R}^{n+1} \to \mathbb{R}$ with $x \mapsto \|x\|^2$, we have $S^n = f^{-1}(\{1\})$. For every $x \neq 0$, $Df(x) \neq 0$, so by implicit function theorem with respect to some co-ordinate system, there exists charts (it turns out an atlas for S^n requires strictly more than 1 chart). Since f is a polynomial (e.g. standard Cartesian co-ordinates), S^n is a C^∞ n-manifold.

Note that 2-sphere would be the standard sphere, which is two-dimensional with zero volume.

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^k$ be a C^1 function. A point $x \in U$ is called a **critical point** if $\operatorname{rank}(Df(x)) < k$, i.e. the columns of the derivative matrix do not span $\mathbb{R}^k$, and $f(x)$ is called a **critical value**. Otherwise x is called a **regular point**.

Example • For $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = \|x\|$, clearly $x = 0$ is the only critical point, and 0 is the associated critical value.

- For $f : \mathbb{R}^n \to \mathbb{R}^k$, if $k > n$ then there are no regular points in $\mathbb{R}^n$ by definition.

- For $f : \mathbb{R}^3 \to \mathbb{R}^2$, is we have $f(x,y,z) = (e^z x, (y-1)\sin z)$, then
$$Df(x,y,z) = \begin{pmatrix} e^z & 0 & e^z x \\ 0 & \sin z & (y-1)\cos z \end{pmatrix}.$$
If $\sin z \neq 0$ then all points are regular since $e^z \neq 0$. If $\sin z = 0$, then $\cos z = \pm 1$, and points with $y \neq 1$ are regular points. Otherwise, the critical points are $(x, 1, n\pi)$ with $n \in \mathbb{Z}$, and the critical values are $f(x, 1, n\pi) = (xe^{n\pi}, 0)$ (or just the whole $y = 0$ line in $\mathbb{R}^2$).

Theorem 2.2.6 *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^k$ be a C^ℓ-map with $\ell \geq 1$, and U is open. If $y \in \mathbb{R}^k$ is a regular value, then $f^{-1}(\{y\})$ is a C^ℓ $(n-k)$-manifold.*

Proof Let $x \in f^{-1}(\{y\})$. Since x is not a critical point, $Df(x)$ has rank k. After rearranging co-ordinates, we can assume that $(\partial f_i / \partial x_j)(x)$ is invertible, with $i = 1, \ldots k$ and $j = n - k + 1$. The existence of the chart follows from the implicit function theorem, and so $f^{-1}(\{y\})$ is a C^ℓ $(n-k)$-manifold by definition. ■

Note that if $y \notin f(U)$ then $\phi = f^{-1}(\{y\})$ is still a $(n-k)$-manifold.

Example • For $f : \mathbb{R}^n \to \mathbb{R}$ with $x \mapsto \|x\|$, we have $S^{n-1} = f^{-1}(\{1\})$ following from previous example.

- For $f(x,y,z) = (e^z x, (y-1)\sin z)$, the inverse of the regular values $f^{-1}(\{(a,b) : b \neq 0\})$ is a 1-manifold.

- For $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^2$ with $(x,y) \mapsto (\|x\|, \|y\|)$, we have $f^{-1}(\{1,1\}) = S^{n-1} \times S^{m-1}$.

3 Tangent spaces and vector fields

Let $M \subset \mathbb{R}^n$ be a C^ℓ k-manifold with $\ell \geq 1$. The **tangent vector** v at $p \in M$ is an element in $\mathbb{R}^n$ of the form $v = \gamma'(0)$ with $\gamma : (-\epsilon, \epsilon) \to M$ being a C^1 curve, and that $\gamma(0) = p$.

The set of all tangent vectors at point $p \in M$ is the **tangent space** $T_p(M)$ at p.

Proposition 3.0.1 *Let $M \subset \mathbb{R}^n$ be a C^ℓ k-manifold, and $p \in M$. Then $T_p(M)$ is a k-vector space of $\mathbb{R}^n$. In fact, if $h : U \to U' \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ is a chart with $h(p) = 0$, then $T_p(M) \subseteq (Dh^{-1}(0))(\mathbb{R}^k \times \{0\})$.*

Proof Let h be a chart, $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p \in U$. We can assume $\gamma : (-\epsilon, \epsilon) \to U \cap M$, so

$$h \circ g : (-\epsilon, \epsilon) \to \mathbb{R}^k \times \{0\},$$

which implies that

$$\gamma = h^{-1} \circ h \circ \gamma,$$

so that

$$v = \gamma'(0) = (Dh^{-1}(h \circ \gamma(0)))(h \circ \gamma)'(0) = Dh^{-1}(0) \cdot \boldsymbol{w},$$

and thus $T_p(M) \subseteq Dh^{-1}(0)(\mathbb{R}^k \times \{0\})$. On the other hand, let $\delta(t) = t\boldsymbol{w}$, $\boldsymbol{w} \in \mathbb{R}^k$, and we get a curve in M via $h^{-1} \circ \delta$. By the chain rule,

$$(h^{-1} \circ \delta)'(0) = Dh^{-1}(0) \cdot \boldsymbol{w},$$

which implies that $Dh^{-1}(0)(\mathbb{R}^k \times \{0\}) \subseteq T_p(M)$, and so $Dh^{-1}(0)(\mathbb{R}^k \times \{0\}) = T_p(M)$.

The chart h is a diffeomorphism so h is injective, which means $\dim(T_p(M)) = \dim(\mathbb{R}^k) = k$, as required. ■

Theorem 3.0.2 *Let $g : U \subset \mathbb{R}^n \to \mathbb{R}^{n-k}$ be a C^ℓ-function, U is open, and $\boldsymbol{c} \in \mathbb{R}^{n-k}$ is a regular value. Then $M = g^{-1}(\{\boldsymbol{c}\})$ is a k-manifold and $T_p(M) = ker\{Dg(p) \ : \ p \in M\}$.* □

Here the kernel is the one induced by the matrix representing the linear map.

Example Let $M = \{(x,y,z) : x^3 + y^3 + z^3 = 1\}$, and $g(x,y,z) = x^3 + y^3 + z^3$. If $p = (1,-1,1)$, then $T_p(M) = \ker(3(1)^2, 3(-1)^2, 3(1)^2) = \ker(3,3,3) = \{(x,y,z) : x+y+z = 0\}$. On the other hand, for $q = (1,0,0)$, we have $T_q(M) = \ker(3,0,0) = \{(x,y,z) : x = 0\}$.

Let $M \subset \mathbb{R}^n$ be a C^1 manifold, $u \subset \mathbb{R}^n$ open, and $M \subset U$ with $f : U \to \mathbb{R}$ a C^1-function. The point $p \in M$ is a **critical point** of $f|_M$ if for every C^1 curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$, $(f \circ \gamma)'(0) = 0$, i.e., the tangent vector is zero at the critical point p.

If $f|_M$ has a local extreme at $p \in M$ then p is a critical point. By the chain rule, $f|_M$ has a critical point exactly when $Df(p)|_{T_p(M)} = 0$.

3.1 Method of Lagrange multipliers

Proposition 3.1.1 *Let $U \subset \mathbb{R}^{n+m}$ be open, $g : U \to \mathbb{R}^n$ be a C^ℓ-function with $\ell \geq 1$, and $0 \in \mathbb{R}^n$ be a regular value of g. For $f : U \to \mathbb{R}$ a C^1-function, $p \in M = g^{-1}(\{0\})$ is a critical point iff there exists some **Lagrange multipliers** $\lambda_1, \ldots \lambda_n \in \mathbb{R}$ with $D(f + \lambda_i g_i)(p) = 0$.*

Einstein summation convention implied.

Proof Assume there exists the relevant Lagrange multipliers, then

$$0 = D(f + \lambda_i g_i)(p) \quad \Leftrightarrow \quad Df(p) = -\lambda_i Dg_i(p).$$

Hence $Df(p)$ is a linear combination of row vectors of $Dg_i(p)$. Note that $Dg_i(p)|_{T_p(M)} = 0$ by the previous theorem, so p is a critical point.

On the other hand, note that $\operatorname{rank}(Dg(p) = n)$ if p is regular, so $Dg_i(p)$ are linear independent row vectors. Note also that $Df(p)$ is a linear map from $\mathbb{R}^{n+m}$ to $\mathbb{R}$, vanishing on $T_p(M)$ which is m-dimensional and sits in the n-dimensional subvector space of the dual space $(R^{n+m})^*$ housing all of the $Dg_i(p)$. Since $Dg_i(p)$ form a basis for this subspace, we must have constants where $Df(p) = -\lambda_i Dg_i(p)$. ∎

The method of Lagrange multipliers gives a method of finding critical points and extrema. Let $F : U \times \mathbb{R}^n \to \mathbb{R}$ with $(x, \lambda_1, \ldots \lambda_n) \mapsto f(x) + \lambda_i g_i(x)$, the previous identity gives

So it is used a lot in optimisation procedures.

$$\frac{\partial F}{\partial x_i} = 0, \qquad \frac{\partial F}{\partial \lambda_j} = 0, \qquad i = 1, \ldots n+m, \quad j = 1, \ldots n. \tag{3.1}$$

Solving the system gives finitely many critical points. Furthermore, if M is compact, then we can find extrema of f via this method.

Example Find the maximum value of $f(x,y) = x + y$ on $M = \{(x,y) : x^4 + y^4 = 1\}$.

Defining $g(x,y) = x^4 + y^4 - 1$, we have $g^{-1}(\{0\}) = M$ and is a manifold. We define

$$F(x,y) = f + \lambda_i g_i = x + y + \lambda(x^4 + y^4 - 1),$$

which results in

$$0 = \frac{\partial F}{\partial x} = 1 + 4\lambda x^3,$$
$$0 = \frac{\partial F}{\partial y} = 1 + 4\lambda y^3,$$
$$0 = \frac{\partial F}{\partial \lambda} = x^4 + y^4 - 1.$$

Since $(0,0) \notin M$, the first two equations give

$$x = y = \left(-\frac{1}{4\lambda}\right)^{1/3},$$

so the constraint results in $\lambda = \pm 8^{1/4}/4$, and the critical points are $\pm(2^{-1/4}, 2^{-1/4})$. The maximum is thus

$$f(2^{-1/4}, 2^{-1/4}) = \frac{2}{\sqrt{2}}.$$

Example Find the extrema of $f(x,y,z) = 5x + y - 3z$ on the intersection of $x + y + z = 0$ with $S^2 = \{(x,y,z) : x^2 + y^2 + z^2 = 1\}$.

Consider

$$F(x,y,z,\lambda,\mu) = 5x + y - 3z + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 1).$$

It can be shown that $\lambda = -1$ from the first three equations. That results in $y\mu = 0$ in the second equation, and for a non-trivial constraint, we thus have $y = 0$. This leads then in $x = -2\mu$, $z = 2/\mu$, resulting in $2x^2 = 1$, and thus the critical points are

$$a = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \quad b = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right).$$

The extrema are then $f(a) = 8/\sqrt{2}$ and $f(b) = -8/\sqrt{2}$.

Proposition 3.1.2 *Let $A \subset \mathbb{R}^n$ be compact, $B \subset \mathbb{R}^n$ be closed, and both non-empty. Then there exists $a \in A$ and $b \in B$ where*

$$\|a - b\| \leq \|x - y\|$$

for all $x \in A$ and $y \in B$, and this can be any norm.

Proof Let $d = \inf\{\|x - y\| : x \in A, y \in B\}$. For all $n \in \mathbb{N}$, there exists some $a_n \in A$ and $b_n \in B$ such that

$$\|a_n - b_n\| < d + \frac{1}{n}.$$

By passing to a sub-sequence, we can assume $a_n \to a$ since A is compact. Then we see that

$$\|b_n\| \leq \|b_n - a_n\| + \|a_n - a\| + \|a\| \leq d + 1 + \|a\|$$

for $n \gg 1$. This implies that $B \cap D(0; d+1+\|a\|)$ is compact, so $b_n \to b$ as $n \to \infty$. Since $b \in B$, we have

$$\|a-b\| \leq \|a-a_n\| + \|a_n - b_n\| + \|b_n - b\| < d + \epsilon$$

for some ϵ. Since d is the infimum, we must have $\|a-b\| \leq \|x+y\|$ for all $x \in A$ and $y \in B$. ∎

Example Find $q \in M = \{x \in \mathbb{R}^3 : 2x^+ y^2 + z = 1\}$ which has minimum distance to $p = (0,0,-5)$.

Now, $M = g^{-1}(\{0\})$ where $g = 2x^2 + y^2 + z - 1$, and since 0 is a regular value, M is closed (but not bounded). Let $f(x,y,z) = x^2 + y^2 + (z+5)^2 = \|x-p\|^2$ be the norm of choice, and minimising the norm gives us the desired solution. Consider

$$F(x,y,z,\lambda) = x^2 + y^2 + (z+5)^2 + \lambda(2x^2 + y^2 + z - 1).$$

The usual manoeuver gives $x = 0$ or $\lambda = -1/2$, which we consider separately.

- For $\lambda = -1/2$, we have $y = 0$, $z = -19/4$, $x = \pm\sqrt{23/8}$, so $f(\pm\sqrt{23/8}, 0, -19/4) = 47/16 < 3$.
- For $x = 0$, we have $y = 0$ or $\lambda = -1$. The former case gives $z = 1$ and thus $f(0,0,1) = 36 > 3$. For $\lambda = -1$, we have $z = -9/2$ and thus $y = \pm\sqrt{11/2}$, which gives $f(0, \pm\sqrt{11/2}, -9/2) = 23/4 > 3$.

So $q = (\pm\sqrt{23/8}, 0, -19/4)$.

3.2 *Tangent spaces*

Let $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ be two C^ℓ manifolds, $\ell \geq 1$. Assume we have a continuous map $f : M \to N$ which extends to a C^1 map $\overline{f} : U \to \mathbb{R}^n$, where $U \supset M$ is open. We define, for $p \in M$,

$$T_p(\overline{f}) : T_p(M) \to T_{\overline{f}(p)}(N),$$

where for $\gamma : (-\epsilon, \epsilon) \to M$ a C^1 curve with $\gamma(0) = p$ and $(\overline{f} \circ \gamma)(0) = \overline{f}(p)$, we have

$$T_p(\overline{f}) = (f \circ \gamma)'(0) \in T_{f(p)}(N).$$

By chain rule,

$$(\overline{f} \circ \gamma)' = D\overline{f}(\gamma(0)) \cdot \gamma'(0) = D\overline{f}(p) \cdot \gamma'(0),$$

which implies that for $T_p(\overline{f}) : T_p(M) \to T_{\overline{f}(p)}(N)$, we have

$$T_p(f(v)) = Df(p) \cdot v.$$

We observe that $v \in T_p(M)$ implies that $D\overline{f}(p) \cdot v \in T_{\overline{f}(p)}(N)$, and $T_p(\overline{f})$ is a linear map between the two tangent spaces.

A map $f : M \to N$ is called a C^ℓ **map** if f extends to a C^ℓ map $\overline{f} : U \to \mathbb{R}^n$ as before. Here, if $T_p(f)$ is not surjective, then $p \in M$ is a **critical point**, and $f(p)$ its **critical value**.

Theorem 3.2.1 *Let $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ be two C^ℓ manifolds, $\ell \geq 1$, and $f : M \to N$ which extends to a C^ℓ map. If $x \in N$ is a regular value, then $f^{-1}(\{x\})$ is a C^ℓ manifold of dimension $\dim(M) - \dim(N)$.*

Proof Let $y \in f^{-1}(\{x\}) \subseteq M$, and we seek a chart around y. Let $s = \dim(M)$ and $r = \dim(N)$. We need

$$\psi : U^r \subset \mathbb{R}^{s-r} \times \mathbb{R}^{m+r-s},$$

with

$$\psi\left(f^{-1}(\{x\} \cap U)\right) = U' \cap \left(\mathbb{R}^{s-r} \times \{0\}\right).$$

Let $g : V \to \mathbb{R}^N \times \mathbb{R}^{n-r}$ be a chart around $x \in N$. Choose U_y such that $y \in U_y$, $f(U_y) \subset V$, and a chart

$$h : U_y \to U_y' \subset \mathbb{R}^s \times \mathbb{R}^{m-s}.$$

We have the following:

$$\begin{tikzcd}
\mathbb{R}^s \times \{0\} \arrow[r] & M \arrow[r, "f"] & N \arrow[r, "\hat{p}"] \arrow[dr, "g"] & \mathbb{R}^r \times \{0\} \\
\mathbb{R}^s \times \mathbb{R}^{m-s} \arrow[ur, "h"] & & & \mathbb{R}^r \times \mathbb{R}^{n-r}
\end{tikzcd}$$

Let $\phi : \mathbb{R}^s \to \mathbb{R}^r$, which has full rank since

- a chart maps tangent plane to tangent plane
- f is surjective by assumption since x is a regular point
- a chart maps to tangent plane.

For $0 \in \mathbb{R}^s$ corresponding to $g \in M$ via h, $D\phi(0)$ has full rank. The same conclusion follows with $0 \in \mathbb{R}^r$ corresponding to $x \in N$ via g. Thus $\phi^{-1}(\{0\})$ is a manifold and corresponds to $f^{-1}(\{x\}) \cap U_y$ by

$$h\left(\phi^{-1}(\{0\} \times \{0\})\right) = f^{-1}(\{x\}) \cap U_y,$$

so ϕ is a chart and $f^{-1}(\{x\})$ is a manifold. ∎

3.3 *Vector fields*

Let $M \subset \mathbb{R}^n$ be a C^ℓ manifold with $\ell \geq 1$. A continuous function $v : M \to \mathbb{R}^n$ is called a **vector field** if $v(x) \in T_x(M)$ for all $x \in M$. It is called a C^ℓ**-vector field** if there is an open set $U \subset \mathbb{R}^n$ containing M such that v extends to a C^ℓ function $\overline{v} : U \to \mathbb{R}^n$.

Example For $S^{n-1} = \{x \in \mathbb{R}^n \ : \ \|x\|^2 = 1\}$. Let $g(x) = \|x\|^2$, then $S^{n-1} = g^{-1}(\{1\})$. By theorem,

$$\begin{aligned} T_p(S^{n-1}) = \ker Dg(p) = \ker(2p) &= \{x \in \mathbb{R}^n \ : \ 2x_i p_i = 0\} \\ &= \{x \in \mathbb{R}^n \ : \ (x, p) = 0\}. \end{aligned}$$

So for a vector field $v(x)$ on S^{n-1}, we need $(x, v(x)) = 0$ for all $x \in S^{n-1}$. For $n = 2m$, let $v : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$, with

$$(x_1, \ldots, x_{2m}) \mapsto (-x_2 x_1, -x_3 x_2, \ldots, -x_{2m} x_{2m-1}).$$

Here we have $(x, v(x)) = 0$ for all $x \in \mathbb{R}^{2m}$, so v restricts to a vector field on S^{2m-1}, is C^∞, and $v(x) \neq 0$ for all $x \in S^{2m-1}$.

v is called a **non-vanishing vector field** in this case. Note that there are no non-vanishing vector fields on S^{2m}.

This is related to the **hairy ball theorem**.

Example For $0 < \epsilon < 1$, define

$$\phi : (1-\epsilon, 1+\epsilon) \times \mathbb{R}^2 \to \mathbb{R}^3; \quad \Phi(r, \phi, \theta) = \begin{pmatrix} (2 + r\cos\phi)\cos\theta \\ (2 + r\cos\phi)\sin\theta \\ r\sin\phi \end{pmatrix},$$

where ϕ and θ are both full angles from 0 to 2π. The 2**-torus** is then

$$T^2 = \{x \in \mathbb{R}^3 \ : \ (x, y, z) = \Phi(1, \phi, \theta)\}.$$

If we restrict Φ to small angles we get charts, and so we get tangent plans and vector fields. Note that the 2-torus does have non-vanishing vector fields, compared to the 2-sphere.

Let $U, V \subset \mathbb{R}^n$ be open sets, $h : U \times V$ be a C^∞ diffeomorphism, and $v : U \to \mathbb{R}^n$ be a vector field. We define the vector field on v by

$$h * v : V \to \mathbb{R}^n, \quad h * v(x) = Dh\left(h^{-1}(x)\right) \cdot v(h^{-1}(x)).$$

Lemma 3.3.1 *If M is a C^∞ manifold and v a C^ℓ vector field, with $\ell \geq 1$. For all $p \in M$, there exists open $I \subset \mathbb{R}$ with $0 \in I$, and an integral curve $\gamma : I \to M$ such that $\gamma(0) = p$, $\gamma'(t) = v(\gamma(t))$ for all $t \in I$.* □

Lemma 3.3.2 *As above, for $i = 1, 2$, let $\gamma_i : I_i \to M$ be integral curves of v with $\gamma_1(0) = p = \gamma_2(0)$, I_i open, and $0 \in I_i$. Then $\gamma_1(t) = \gamma_2(t)$ for all $t \in I_1 \cap I_2$.*

Proof Uniqueness follows from the Picard–Lindelöf theorem. ■

Note that the integral curve can now be extended to an integral curve of $I_1 \cup I_2$, and we get a maximal curve through a point p this way.

Proposition 3.3.3 *Let M be a compact C^∞ manifold and v and C^ℓ vector field with $\ell \geq 1$. For all $p \in M$, there exists $\gamma : \mathbb{R} \to M$ with $\gamma(0) = p$.*

Proof Let $\gamma : I \to M$ be the maximal integral curve, and assume $I \cap [0, \infty)$ is bounded. Then these exists $T = \sup\{I \cap [0, \infty)\}$. Choose a sequence $(t_n) \in I$ with $t_n \to T$, then $\gamma(t_n)$ is a sequence in M. Since M is compact, we can assume $\gamma(t_n) \to x \in M$.

Let $\beta : (T - \epsilon, T + \epsilon) \to M$ be an integral curve with $\beta(T) = x$. Since $t_n \to T$ for large n, and $t_n \in (T - \epsilon, T + \epsilon)$, we should have $\gamma(t_n) = \beta(t_n)$ by uniqueness, and so γ can be extended beyond T. However, this is a contradiction since γ was assumed to be maximal, so I is not bounded, and thus γ can be extended to $\mathbb{R}$. ∎

For v a C^ℓ vector field ($\ell \geq 1$) on a compact manifold M, the **flow** Φ is defined as

$$\Phi : M \times \mathbb{R} \to M, \qquad (x, t) \mapsto \gamma_x(t) \tag{3.2}$$

where γ_x is the integral curve with $\gamma_x(0) = x$.

Theorem 3.3.4 *Let M be a compact C^∞ manifold and v a C^ℓ vector field, $\ell \geq 1$. Then the flow Φ is continuous and*

1. *$\Phi(x, 0) = x$ for all $x \in M$,*
2. *$\Phi(\Phi(x, t), x) = \Phi(x, t + s)$ for all $x \in M$, $t, s \in \mathbb{R}$.*

Proof Continuity holds and follows from Picard–Lindelöf, and $\Phi(x, 0) = x$ follows from definition of the flow map. Let $y = \Phi(x, t)$, so $\gamma_x(t) = y$. Define $\gamma(u) = \gamma_x(u + t)$, which is an integral curve with $\gamma(0) = y$. By uniqueness, $\gamma = \gamma_y$, and so

$$\Phi(\Phi(x, t), s) = \gamma_y(s) = \gamma(s) = \gamma_x(s + t) = \Phi(x, t + s).$$

∎

Note that if we write $x \cdot t = \Phi(x, t)$, then $x \cdot 0 = x$, and $(x \cdot t) \cdot s = x \cdot (t + s)$, so the abelian group $\mathbb{R}$ acts on the set M. Since Φ is continuous we have a topological action. Every C^1 vector field v on a compact manifold M gives rise to an $\mathbb{R}$-action on M.

Note also that v is of C^ℓ class implies that Φ is of C^ℓ class.

4 Differential forms on $\mathbb{R}^n$

In lower dimensions, from the standard fundamental theorem of calculus, Stokes' theorem and divergence theorem, we see we have identities of the form

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega, \tag{4.1}$$

where M is some (oriented) manifold, and ω is some function / vector field. This is in fact true in higher dimensions, and the result is the **generalised Stokes' theorem**. It will be seen ω is a **differential k-form**, and M are the **oriented ℓ-manifolds** in $\mathbb{R}^n$ with boundary ∂M. To get to the general result, we go through some machinery first in $\mathbb{R}^n$, before proceeding to general (oriented) manifolds.

4.1 Riemann integrals

For $f : [a,b] \to \mathbb{R}$, recall that for a partition $Z = \{t_0, t_1, \ldots, t_n\}$, the **upper/lower Riemann sums** are defined as

$$\mathcal{U}(f,Z) = \sum_{i=0}^{n-1} M_i(f)(t_{i+1} - t_i), \qquad \mathcal{L}(f,Z) = \sum_{i=0}^{n-1} m_i(f)(t_{i+1} - t_i), \tag{4.2}$$

where for $x \in [t_{i-1}, t_i]$,

$$M_i(f) = \sup f(x), \quad m_i(f) = \inf f(x).$$

If Z' is a **refinement** of Z (i.e. $Z' \supset Z$, where Z' is a partition), then

$$\mathcal{L}(f,Z) \leq \mathcal{L}(f,Z') \leq \mathcal{U}(f,Z') \leq \mathcal{U}(f,Z).$$

For two partitions, the **common refinement** is $Z'' = Z' \cup Z$, which implies that

$$\mathcal{L}(f,Z) \leq \mathcal{L}(f,Z'') \leq \mathcal{U}(f,Z'') \leq \mathcal{U}(f,Z').$$

The **upper Riemann integral** is then defined as

$$\int_{[a,b]}^{u} f \,\mathrm{d}x = \inf\{\mathcal{U}(f,Z) \ : \ Z \text{ a partition of } [a,b]\}, \tag{4.3}$$

while the **lower Riemann integral** is

$$\int_{[a,b]}^{l} f \, \mathrm{d}x = \inf\{\mathcal{L}(f, Z) \ : \ Z \text{ a partition of } [a, b]\}. \tag{4.4}$$

For bounded f, we should have

$$\int_{[a,b]}^{l} f \, \mathrm{d}x \le \int_{[a,b]}^{u} f \, \mathrm{d}x \le \infty.$$

If the two sums coincide as $|t_{i-1} - t_i| \to 0$, then f is **Riemann integrable**.

In $\mathbb{R}^n$, to generalise, f defined analogously if each individual component of f is Riemann integrable.

Lemma 4.1.1 *Let $f : [a, b] \to \mathbb{R}^n$ be integrable. Then $\|f\|$ is also integrable and*

$$\left\| \int_{[a,b]} f \, \mathrm{d}x \right\| \le \int_{[a,b]} \|f\| \, \mathrm{d}x.$$

Proof $\|f\|$ is clear integrable. We see that for all $\epsilon > 0$, there exists a common partition Z such that

$$\mathcal{U}(f_i, Z) - \mathcal{L}(f_i, Z) \le \epsilon, \quad \mathcal{U}(\|f\|, Z) - \mathcal{L}(\|f\|, Z) \le \epsilon$$

for all components f_i of f. For any partition $Z = \{x_0, \ldots, x_n\}$ and any choice ξ_i

$$a = x_0 \le \xi_0 \le x_1 \le \xi_1 \le \ldots \le x_{n-1} \le \xi_{n-1} \le x_n = b,$$

we have

$$\left\| \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i) \right\| \le \sum_{i=0}^{n-1} \|f(\xi_i)\|(x_{i+1} - x_i)$$

by the triangle inequality. For such a partition Z, we have both

$$\left| \int_a^b f_i \, \mathrm{d}x - \sum_{k=0}^{n-1} f_i(\xi_k)(x_{k+1} - x_k) \right| \le \epsilon,$$

$$\left| \int_a^b \|f\| \, \mathrm{d}x - \sum_{k=0}^{n-1} \|f(\xi_k)\|(x_{k+1} - x_k) \right| \le \epsilon,$$

which implies that, considering each component,

$$\left\| \int_a^b f \, \mathrm{d}x - \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k) \right\| \le \sqrt{n\epsilon^2} = \sqrt{n}\epsilon.$$

Then,

$$\begin{aligned} \left\| \int_{[a,b]} f \, \mathrm{d}x \right\| &\le \left\| \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k) \right\| + \sqrt{n}\epsilon \\ &\le \sum_{k=0}^{n-1} \|f(\xi_k)(x_{k+1} - x_k)\| + \sqrt{n}\epsilon \\ &\le \int_a^b \|f\| \, \mathrm{d}x + \epsilon + \sqrt{n}\epsilon. \end{aligned}$$

Since ϵ was arbitrary, we have the result as required. ■

4.2 *Differential 1-forms and line integrals*

Let V be a real vector space with norm $\|\cdot\|$, and $c : [a,b] \to V$ be continuous. The **length** of c is defined as

$$L(c) = \sup \left\{ \sum_{i}^{n-1} \|c(t_{i+1} - c(t_i)\| \ : \ \forall\, n \in N, t_i\, inZ \right\}.$$

The curve c is **rectifiable** if $L(c) < \infty$.

Note that the length of a curve is (and should be) independent of its parameterisation.

Proposition 4.2.1 *For $c : [a,b] \to^n$ of class C^1, c is rectifiable, and $L(c) = \int_a^b \|c'(t)\|\, dt$.*

Proof Note that

$$\begin{aligned}
\sum_i \|c(t_{i+1} - c(t_i)\| &= \sum_i \left\| \int_{[t_i,t_{i+1}]} c'(t)\, dt \right\| \\
&\leq \sum_i \int_{[t_i,t_{i+1}]} \|c'(t)\|\, dt \\
&= \int_a^b \|c'(t)\|\, dt < \infty
\end{aligned}$$

since $c \in C^1[a,b]$, so $L(c) < \infty$ and c is rectifiable.

Let $f(t) = L\left(c|_{[a,t]}\right)$ for $a \leq t_0 < t \leq b$. Then

$$\begin{aligned}
\left| \frac{c(t) - c(t_0)}{t - t_0} \right| \leq \frac{L\left(c|_{[t_0,t]}\right)}{t - t_0} &= \frac{f(t) - f(t_0)}{t - t_0} \\
&\leq \frac{1}{t - t_0} \int_{[t_0,t]} \|c'(s)\|\, ds = \|c'(t_1)\|
\end{aligned}$$

for $t_1 \in [t_0, t]$ by the mean value theorem. So as $t_0, t_1 \to t$,

$$\|c'(t)\| \leq f'(t) \leq \|c'(t)\|,$$

and so $f'(t)$ exists and $f'(t) = \|c'(t)\|$. Thus

$$L(c) = f(b) - f(a) = \int_a^b f'(t)\, dt = \int_a^b \|c'(t)\|\, dt < \infty$$

as required. ■

Example For the helix, we have $c(t) = (at, r\cos t, r\sin t)$, so $\|c'(t)\| = \sqrt{r^2 + a^2}$, and

$$L(c)|_{[0,2\pi]} = \int_0^{2\pi} \sqrt{r^2 + a^2}\, dt = 2\pi\sqrt{r^2 + a^2}.$$

Recall the differential $Df(x) : \mathbb{R}^n \to \mathbb{R}$ is a linear map. Denote this now has $\mathrm{d}f_x : \mathbb{R}^n \to \mathbb{R}$, and the real vector space of all linear maps $\phi : \mathbb{R}^n \to \mathbb{R}$ to be $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$. Since $\phi \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ is linear, we only investigate the action of ϕ on the basis $\{e_i\}$. We see that

$$\mathrm{d}f_x(e_i) = \frac{\partial f}{\partial x_i}(x),$$

so

Einstein summation.

$$\mathrm{d}f_x(v) = \frac{\partial f}{\partial x_i} v_i = \langle \nabla f, v \rangle$$

for $\langle \cdot, \cdot \rangle$ the inner product on some vector space V housing v. We see $\partial f / \partial x_i$ are smooth coefficient functions, and that since $f : U \to \mathbb{R}^n$ is smooth, $\mathrm{d}f : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

Let $U \subset \mathbb{R}^n$ be open and $\omega : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, then there are functions $f_1, \ldots f_n : U \to \mathbb{R}^n$ where

$$\omega_x(v) = f_i(x) v_i \tag{4.5}$$

for all $x \in U$, $v = v_i e_i$. The coefficient functions f_i are calculated via

$$f_i(x) = \omega_x(e_i). \tag{4.6}$$

We call ω a **differential 1-form** on U if $f_i : U \to \mathbb{R}$ are of class C^∞ for all i. The set of all 1-forms is denoted by $\Omega^1(U)$, which has the structure of a real vector space. One can canonically multiply a 1-form $\omega \in \Omega^1(U)$ with a smooth function $f \in C^\infty(U)$ by performing

$$(f\omega)_x(v) = f(x)\omega_x(v).$$

Lemma 4.2.2 *Let ω be a 1-form and $U \subset \mathbb{R}^N$ be open. Then there exists a smooth vector field $X_\omega : U \to \mathbb{R}^n$ where*

$$\omega_x(v) = \langle X_\omega(x), v \rangle.$$

Proof Since $F_\omega = f_i e_i$, $\langle F_\omega(x), v \rangle = f_i(x) \langle e_i, v \rangle = f_i(x) v_i = \omega_x(v)$. ■

Lemma 4.2.3 *Let $\omega \in \Omega^1(U)$, $x_i : U \to \mathbb{R}$, $x_i(p_1, \ldots, p_n) = p_i$. Then $\omega = f_i \, \mathrm{d}x_i$ and $f_i(x) = \omega_x(e_i)$.*

Proof Since

$$\mathrm{d}x_i(x)(v) = \frac{\partial x_i}{\partial x_j}(x) v_j = v_i,$$

we have

$$\omega_x(v) = f_i(x) v_i = f_i(x) \, \mathrm{d}x_i(x)(v),$$

and since v was arbitrary, $\omega_x = f_i(x) \, \mathrm{d}x_i$. ■

Example Suppose $\omega \in \Omega^1(\mathbb{R}^2)$ and $\omega = 3xy\ \mathrm{d}x + y^3\ \mathrm{d}y$, and we take $p = (7,3)$. Then since

$$\omega_p(e_1) = 3 \cdot p_1 \cdot p_2\ \mathrm{d}x(e_1) + p_2^3\ \mathrm{d}y(e_1) = 3 \cdot 7 \cdot 1 + 27 \cdot 0 = 63,$$

while $\omega_p(e_2) = 27$ by a similar argument. So $\omega_p((1,-2)) = 63 - 2 \cdot 27 = 9$ for example.

A 1-form eats a vector and spits out a number. Sometimes it can be regarded as a functional (eats a function and spits out a number).

A differential 1-form ω is **exact** if there exists some $f \in C^\infty(U)$ where

$$\omega = \mathrm{d}f. \tag{4.7}$$

For $U \subset \mathbb{R}^n$ open and $c : [a,b] \to U$ be smooth and $\omega \in \Omega^1(U)$. The **line integral** of ω along c is

$$\int_c \omega = \int_a^b \omega_{c(t)}\left(c'(t)\right)\ \mathrm{d}t. \tag{4.8}$$

If c is piecewise smooth, we can still define the integral.

For $c : [a,b] \to \mathbb{R}^n$ a smooth curve and $\phi : [\alpha,\beta] \to [a,b]$ be a smooth bijective map, then $\tilde{c} = c \circ \phi : [\alpha,\beta] \to \mathbb{R}^n$ is a **orientation preserving reparameterisation** if $\phi' > 0$, otherwise it is orientation reversing.

Proposition 4.2.4 *For $\omega \in \Omega^1(U)$, $c : [a,b] \to U$, $\tilde{c} = c \circ \phi : [\alpha,\beta] \to U$ orientation preserving,*[1]

$$\int_c \omega = \int_{\tilde{c}} \omega.$$

[1] If orientation reversing, then there would be an extra minus sign.

Proof

$$\begin{aligned}\int_{\tilde{c}} \omega = \int_\alpha^\beta \omega_{\tilde{c}}(\tilde{c}')\ \mathrm{d}t &= \int_\alpha^\beta \omega_{c\circ\phi}\left((c \circ \phi)'\right)\ \mathrm{d}t \\ &= \int_\alpha^\beta \omega_{c\circ\phi}\left((c' \circ \phi)\right)\phi'\ \mathrm{d}t \\ &= \int_a^b \omega_c\left(c'\right)\ \mathrm{d}t = \int_c \omega.\end{aligned}$$

■

Lemma 4.2.5 *For $f \in C^\infty(U)$, $c : [a,b] \to U$, then*

$$\int_c \mathrm{d}f = f(c(b)) - f(c(a)).$$

Proof

$$\int_c \mathrm{d}f = \int_a^b \left[Df(c(t))\right]\left(c'(t)\right)\mathrm{d}t = \int_a^b (f \circ c)'(t)\ \mathrm{d}t = f(c(b)) - f(c(a)).$$

■

Proposition 4.2.6 *Let $U \subset \mathbb{R}^n$ be open and path connected, then the following are equilvalent:*

1. *$\omega \in \Omega^1(U)$ is exact,*
2. *$\int_c \omega$ depends only on the end points (i.e. path independence),*
3. *$\oint_c \omega = 0$.*

Proof • 1) implies 2) by previous lemma.

- 2) implies 3). Consider c with $c(a) = c(b) = p \in U$, and $c_p(t) = p$ for all t. Then since $c_p'(t) = 0$, $\int_{c_p} \omega = 0$ and so $\oint_c \omega = 0$ since we assumed path independence.
- 3) implies 2). For $c = c_1 \cup (-c_2)$,
$$0 = \oint_c \omega = \int_{c_1} \omega + \int_{-c_2} \omega = \int_{c_1} \omega - \int_{c_2} \omega,$$
which implies path independence.
- 2) implies 1). Choosing $p \in U$, $f : U \to \mathbb{R}$ via $f(q) = \int_c \omega$, $c(0) = p$, $c(1) = q$, then f is well-defined by 2). Choose $h \in \mathbb{R}^n$ where $h + q \in U$, let
$$\gamma : [0,1] \to I, \quad \gamma(t) = q + th.$$
Then
$$f(q+th) - f(q) = \int_\gamma \omega = \int_0^1 \omega_\gamma(\gamma')\,\mathrm{d}t \int_0^1 \omega_{q+th}(h)\,\mathrm{d}t.$$
Now introduce vector field X_ω, an by lemma,
$$\begin{aligned} f(q+th) - f(q) - \omega_q(h) &= \int_0^1 \Big(\omega_{q+th}(h) - \omega_q(h)\Big)\,\mathrm{d}t \\ &= \int_0^1 \langle X_\omega(q+th) - X_\omega(q), h\rangle\,\mathrm{d}t \\ &\le \int_0^1 \|X_\omega(q+th) - X_\omega(q)\|\,\mathrm{d}t \cdot \|h\|. \end{aligned}$$
The right hand side has $R(h)/\|h\| \to 0$ as $h \to 0$ since X_ω is continuous, so f is differentiable, and
$$\omega_q(h) = Df(q)(h) = \mathrm{d}f_q(h),$$
and since q and h are arbitrary, $\omega = \mathrm{d}f$.

■

Proposition 4.2.7 *If $\omega \in \Omega^1(U)$ is exact and $\omega = f_i\,\mathrm{d}x_i$, then*
$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$$
for all i and j.

Proof Since

$$\omega = \mathrm{d}f = \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i,$$

we have $f_i = \partial f / \partial x_i$, so then since f is differentiable,

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f_j}{\partial x_i}.$$

■

$\omega \in \Omega^1(U)$ is **closed** if for $\omega = f_i\,\mathrm{d}x_i$, $\partial f_i / \partial x_j = \partial f_j / \partial x_i$.

4.3 *Differential k-forms*

4.4 *Integration in* $\mathbb{R}^n$

5 Differential forms on oriented manifolds

5.1 Stokes' theorem