Mathematical Studies Analysis

Notes taken by Runqiu Ye
Lectures by Ian Tice
Carnegie Mellon University

Spring 2025

Contents

1 Advanced topics in metric space theory

1.1 Baire category

Definition. Let X be a metric space.

1. We say that $E \subset X$ is nowhere dense if $(\overline{E})^\circ = \emptyset$.
2. We say that $E \subset X$ is meager in X if

$$E = \bigcup_{\alpha \in A} E_\alpha,$$

where A is a countable set and $E_\alpha \subset X$ is nowhere dense for every $\alpha \in A$.

Theorem. Prove that the following are equivalent for $E \subset X$:

1. E is nowhere dense
2. $\overline{E}$ is nowhere dense
3. $(\overline{E})^c$ is open and dense in X.

Proof. (1) $\implies$ (2). Suppose E is nowhere dense, then $(\overline{E})^\circ = \emptyset$. Note that the closure of $\overline{E}$ is just $\overline{E}$ itself. It follows that $\overline{E}$ is also nowhere dense.

(2) $\implies$ (3). Suppose $\overline{E}$ is nowhere dense. Note that $\overline{E}$ is closed, so $(\overline{E})^c$ is open. Let $x \in X$ be arbitrary. Since $\overline{E}$ is nowhere dense, $x \notin (\overline{E})^\circ$. This implies that for arbitrary $\varepsilon > 0$, we have $B(x, \varepsilon) \not\subset \overline{E}$. This is equivalent to $B(x, \varepsilon) \cap (\overline{E})^c \neq \emptyset$. Hence, $(\overline{E})^c$ is dense in X.

(3) $\implies$ (1). Suppose $(\overline{E})^c$ is dense in X. Let $x \in X$ and $\varepsilon > 0$ be arbitrary. It follows that $B(x, \varepsilon) \cap (\overline{E})^c \neq \emptyset$. This is equivalent to $B(x, \varepsilon) \not\subset \overline{E}$. Therefore, $(\overline{E})^\circ = \emptyset$ and E is nowhere dense.

□

Theorem (Baire category theorem)**.** Let X be a complete metric space. Suppose that for each $n \in \mathbb{N}$, $U_n \subset X$ is open and dense in X. Prove that $\bigcap_{n=0}^{\infty} U_n$ is dense in X. Hint: use the shrinking closed set property.

Proof. Consider any $x \in X$ and arbitrary $\varepsilon > 0$, it suffices to show that $U_n \cap B(x, \varepsilon) \neq \emptyset$ for each $n \in \mathbb{N}$. Now inductively choosing a sequence $x_i \in X$ and $\varepsilon_i > 0$ such that for each $i \in \mathbb{N}$, $B[x_i, \varepsilon_i] \subset U_i$, $B[x_{i+1}, \varepsilon_i] \subset B[x_i, \varepsilon_i] \subset B(x, \varepsilon)$, and $\varepsilon_i < 2^{-i}\varepsilon$.

Since U_0 is dense in X, $B(x, \varepsilon) \cap U_0 \neq \emptyset$. Note that both U_0 and $B(x, \varepsilon)$ are open, so we can choose $x_0 \in B(x, \varepsilon) \cap U_0$ and $\varepsilon_0 > 0$ so small that $B[x_0, \varepsilon_0] \subset B(x, \varepsilon) \cap U_0$ and $\varepsilon_0 < \varepsilon$. Now suppose for $0 \leq i \leq n$, we have chosen $x_i \in X$ and $\varepsilon_i > 0$ such that $B[x_i, \varepsilon_i] \subset U_i$ and $\varepsilon_i < 2^{-i}\varepsilon$ for all $0 \leq i \leq n$, $B[x_{i+1}, \varepsilon_{i+1}] \subset B[x_i, \varepsilon_i]$ for all $0 \leq i < n$. Since U_{n+1} is dense in X, $B(x_n, \varepsilon_n) \cap U_{n+1} \neq \emptyset$. Note also both U_{n+1} and $B(x_n, \varepsilon_n)$ are open. Therefore, choose $x_{n+1} \in B(x_n, \varepsilon_n) \cap U_{n+1}$ and $\varepsilon_{n+1} > 0$ so small that $B[x_{n+1}, \varepsilon_{n+1}] \subset B(x_n, \varepsilon_n) \cap U_{n+1}$ and

$\varepsilon_{n+1} < \frac{\varepsilon_n}{2}$. It follows that $B[x_{n+1}, \varepsilon_{n+1}] \subset U_{n+1}$ and $B[x_{n+1}, \varepsilon_{n+1}] \subset B[x_n, \varepsilon_n] \subset B(x, \varepsilon)$. Also, $\varepsilon < \frac{\varepsilon_n}{2} < 2^{-n-1}\varepsilon$. Now we have successfully constructing the desired sequence.

Since X is complete, $\bigcap_{n=0}^{\infty} B[x_n, \varepsilon_n] = \{z\}$ for some $z \in X$. Note that for each n, we have $z \in B[x_n, \varepsilon_n] \subset U_n$. Also, $z \in B[x_n, \varepsilon_n] \subset B(x, \varepsilon)$. Therefore, $z \in U_n \cap B(x, \varepsilon)$ for each $n \in \mathbb{N}$ and $\bigcap_{n=0}^{\infty} U_n$ is dense in X.

□

Remark. An equivalent statement of the theorem is the following:
Let X be a complete metric space and $\{C_n\}$ a countable collection of closed subsets of X such that $X = \bigcup_{n\in\mathbb{N}} C_n$. Then at least one of the C_n contains an open ball.

1.2 Open mapping theorem

Linear surjections

Theorem (Open mapping theorem)**.** Let X, Y be Banach spaces over a common field and assume that $T \in \mathcal{L}(X; Y)$. Prove that the following are equivalent.

1. T is surjective.
2. There exists $\delta > 0$ such that $B_Y(0, \delta) \subset \overline{T(B_X(0,1))}$.
3. For every $\varepsilon > 0$ there exists $\delta > 0$ such that $B_Y(0, \delta) \subset T(B_X(0, \varepsilon))$.
4. T is an open map: if $U \subset X$ is open, then $T(U) \subset Y$ is open.
5. There exists $C \geq 0$ such that for each $y \in Y$ there exists $x \in X$ such that $Tx = y$ and

$$\|x\|_X \leq C \|y\|_Y .$$

HINT: Prove that (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) $\Longrightarrow$ (5) $\Longrightarrow$ (1), keeping in mind the following suggestions.

1. For (1) $\Longrightarrow$ (2): Study the sets $C_n = \overline{T(B_X(0,n))} \subset Y$ for $n \geq 1$.
2. For (2) $\Longrightarrow$ (3): Prove that $\overline{T(B_X(0,1))} \subset T(B_X(0,3))$ by considering $y \in \overline{T(B_X(0,1))}$ and inductively constructing $\{x_j\}_{j=0}^{\infty} \subset X$ such that $\|x_j\|_X < 2^{-j}$ and $y - \sum_{j=0}^{m} Tx_j \in B_Y(0, 2^{-m-1}R)$ for all $m \in \mathbb{N}$.

Proof. (1) $\Longrightarrow$ (2). Following the hint, for $n \geq 1$ let $C_n = \overline{T(B_X(0,n))}$. Then each of the C_n are closed. Since T is surjective, $Y = \bigcup_{n=1}^{\infty} C_n$. Suppose for contradiction that each C_n are nowhere dense. It then follows that C_n^c are dense in Y. By Baire Category Theorem, $\bigcap_{n=1}^{\infty} C_n^c$ is dense in Y. However, $\bigcap_{n=1}^{\infty} C_n^c = \left(\bigcup_{n=1}^{\infty} C_n\right)^c = \emptyset$, a contradiction. Therefore, at least one C_n is not nowhere dense. That is, there exists some $n \geq 1$, $\overline{T(B_X(0,n))}$ contains an open ball. However, this is the same set as $n\overline{T(B_X(0,1))}$. Therefore, $\overline{T(B_X(0,1))}$ contains an open ball $B_Y(y_0, 4r)$ for some $y_0 \in Y$ and $r > 0$.

Let $y_1 = Tx_1$ for some $x_1 \in B_Y(0,1)$ such that $\|y_0 - y_1\| < 2r$. It follows that $B_Y(y_1, 2r) \subset B_Y(y_0, 4r) \subset \overline{T(B_X(0,1))}$. For any $y \in Y$ such that $\|y\| < r$, we have

$$y = -\frac{1}{2}y_1 + \frac{1}{2}(2y + y_1) = -T\left(\frac{x_1}{2}\right) + \frac{1}{2}(2y + y_1).$$

However, notice that

$$\frac{1}{2}(2y + y_1) \subset \frac{1}{2}B_Y(y_1, 2r) \subset \frac{1}{2}\overline{T(B_X(0,1))} = \overline{T(B_X(0, \tfrac{1}{2}))}.$$

It follows that

$$y = -T\left(\frac{x_1}{2}\right) + \frac{1}{2}(2y + y_1) \in -T\left(\frac{x_1}{2}\right) + \overline{T(B_X(0, \tfrac{1}{2}))}.$$

Note that $-T(\frac{x_1}{2}) \in T(B_X(0, \frac{1}{2}))$. Therefore, $y \in \overline{T(B_X(0,1))}$. Since y is arbitrary with $\|y\| < r$, we have $B_Y(0, r) \subset \overline{T(B_X(0,1))}$.

(2) $\implies$ (3). Following the hint, we first show $\overline{T(B_X(0,1))} \subset T(B_X(0,3))$. By assumption, we have $B_Y(0, R) \subset \overline{T(B_X(0,1))}$ for some $R > 0$. It follows from homogeneity that for each $m \in \mathbb{N}$, we have

$$2^{-m}B_Y(0, R) = B_Y(0, 2^{-m}R) \subset 2^{-m}\overline{T(B_X(0,1))} = \overline{T(B_X(0, 2^{-m}))}.$$

Let $y \in \overline{T(B_X(0,1))}$ and pick $x_0 \in X$ with $\|x\| < 1$ such that $\|y - Tx\| < 2^{-1}R$. Now suppose we have chosen x_j for $0 \le j \le m$ such that $\|x_j\| < 2^{-j}$ and $y - \sum_{j=0}^{m} Tx_j \in B_Y(0, 2^{-m-1}R)$ for all $m \in \mathbb{N}$. By the inclusion above, we can pick $x_{m+1} \in X$ with $\|x_{m+1}\| < 2^{-m-1}$ such that

$$\left\| y - \sum_{j=0}^{m} Tx_j - Tx_{m+1} \right\| = \left\| y - \sum_{j=0}^{m+1} Tx_j \right\| < 2^{-m-2}R.$$

Therefore, $y - \sum_{j=0}^{m+1} Tx_j \in B_Y(0, 2^{-m-2})R$. This completes the inductive construction, and we have found a sequence $\{x_j\}$ such that $\|x_j\| < 2^{-j}$ and $y - \sum_{j=0}^{m} Tx_j \in B_Y(0, 2^{-m-1}R)$ for each $m \in \mathbb{N}$. Note that

$$\sum_{j=0}^{\infty} \|x_j\| \le \sum_{j=0}^{\infty} 2^{-j} = 2,$$

so $\sum_{j=0}^{\infty} x_j$ converges absolutely. Since X is Banach, $\sum_{j=0}^{\infty} x_j$ converges to some $x \in X$ with $\|x\| \le 2$. Also, since $y - \sum_{j=0}^{m} Tx_j \in B_Y(0, 2^{-m-1}R)$, taking the limit where m approaches infinity we obtain

$$y = \sum_{j=0}^{\infty} Tx_j = T\left(\sum_{j=0}^{\infty} x_j\right) = Tx.$$

Therefore, $y \in T(B_X(0,3))$ and thus $\overline{T(B_X(0,1))} \subset T(B_X(0,3))$.

Now for every $\varepsilon > 0$, we have $\frac{\varepsilon}{3}\overline{T(B_X(0,1))} \subset \frac{\varepsilon}{3}T(B_X(0,3)) = T(B_X(0,\varepsilon))$. By assumption, there exists $\delta > 0$ such that $B_Y(0, \delta) \subset \overline{T(B_X(0,1))}$. Therefore,

$$B_Y\left(0, \frac{\delta\varepsilon}{3}\right) = \frac{\varepsilon}{3}B_Y(0, \delta) \subset \frac{\varepsilon}{3}\overline{T(B_X(0,1))} \subset T(B_X(0,\varepsilon)).$$

(3) $\implies$ (4). Let $U \subset X$ be open and $y \in T(U)$. There exists $x \in U$ such that $Tx = y$. Since U is open, there exists $\varepsilon > 0$ such that $B_X(x, \varepsilon) \subset U$. By assumption, there exists $\delta > 0$ such that $B_Y(0, \delta) \subset T(B_X(0, \varepsilon))$. It follows that

$$B_Y(y, \delta) = y + B_Y(0, \delta) \subset Tx + T(B_X(0, \varepsilon)) = T(x + B_X(0, \varepsilon)) \subset T(U).$$

Therefore, $T(U)$ is open and T is an open map.

(4) $\implies$ (5). Since T is an open map, $T(B_X(0, 1))$ is open. Also, $T(0) = 0$ so there exists $r > 0$ such that $B_Y(0, r) \subset T(B_X(0, 1))$. Now let $y \in Y$. Then, $\frac{r}{2\|y\|} y \in B_Y(0, r)$ and there exists $x \in B_X(0, 1)$ such that $Tx = \frac{r}{2\|y\|} y$. It follows that

$$T\left(\frac{2\|y\|}{r} x\right) = y,$$

and since $x \in B_X(0, 1)$,

$$\left\|\frac{2\|y\|}{r} x\right\| = \frac{2\|y\|\|x\|}{r} < \frac{2}{r}\|y\|.$$

Letting $C = \frac{2}{r}$ completes the proof.

(5) $\implies$ (1). Since for each $y \in Y$ there exists $x \in X$ such that $Tx = y$, T is surjective.

□

Linear homeomorphisms, norm equivalence, and closed graphs

Theorem. Let X and Y be Banach spaces and suppose that $T \in \mathcal{L}(X, Y)$ is a bijection. Prove that $T^{-1} \in \mathcal{L}(Y, X)$, and in particular T is a linear (and thus bi-Lipschitz) homeomorphism.

Proof. Since $T \in \mathcal{L}(X, Y)$ is a bijection, T is a surjection. It follows that T is an open map. In particular, for any $U \subset X$ open, $T(U) = (T^{-1})^{-1}(U)$ is open. Therfore, T^{-1} is continuous and thus T is a linear homeomorphism.

□

Theorem. Let X be a vector space that is complete when equipped with both of the norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Prove that if there exists a constant $C_1 > 0$ such that $\|x\|_2 \leq C_1\|x\|_1$ for all $x \in X$, then there exists a constant $C_0 > 0$ such that $C_0\|x\|_1 \leq \|x\|_2 \leq C_1\|x\|_1$ for all $x \in X$.

Proof. Let $T : X_1 \to X_2$, where X_1 and X_2 are X equipped with norms $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively, be the identity map. Then for any $x \in X$ with $\|x\|_1 = 1$, we have

$$\|Tx\|_2 = \|x\|_2 \leq C_1\|x\|_1 = C_1.$$

Therefore, $T \in \mathcal{L}(X_1, X_2)$. T is also surjective. Therefore, there exists a constant $C \geq 0$ such that each $\|x\|_1 \leq C\|x\|_2$. Hence, for each $x \in X$

$$\frac{1}{C}\|x\|_1 \leq \|x\|_2 \leq C_1\|x\|_1.$$

Letting $C_0 = \frac{1}{C}$ completes the proof.

□

Theorem. Let X and Y be Banach spaces and let $T : X \to Y$ be linear (just the algebraic condition). Prove that the following are equivalent

1. T is continuous, i.e. $T \in \mathcal{L}(X; Y)$.
2. The graph of T, $\Gamma(T) = \{(x, Tx) : x \in X\} \subset X \times Y$, is closed in $X \times Y$, where $X \times Y$ is endowed with any of the usual p-norms.

Proof. (a) $\implies$ (b). Let $\{(x_n, Tx_n)\}$ be a convergent sequence in $\Gamma(T)$. Since X is Banach, $x_n \to x$ for some $x \in X$. Since $T \in \mathcal{L}(X; Y)$, we have

$$\lim_{n\to\infty} Tx_n = T\left(\lim_{n\to\infty} x_n\right) = Tx.$$

Therefore, $(x_n, Tx_n) \to (x, Tx) \in \Gamma(T)$, and thus $\Gamma(T)$ is closed.

(b) $\implies$ (a). Let $\pi_1 : \Gamma(T) \to X$ and $\pi_2 : \Gamma(T) \to Y$ by $\pi_1(x, Tx) = x$ and $\pi_2(x, Tx) = Tx$. Since $\Gamma(T)$ is a closed in Banach space Y, $\Gamma(T)$ is Banach space. It is clear that both π_1 and π_2 are bounded linear maps. Moreover, π_1 is a bijection. It follows that $S = \pi_1^{-1}$ is a bounded linear map. Therefore, $T = \pi_2 \circ S$ is a bounded linear map.

□

Linear injections with closed range

Theorem. Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Prove the following are equivalent.

1. T is injective and range(T) is closed.
2. $T : X \to \text{range}(T)$ is a linear homeomorphism.
3. There exists $C \geq 0$ such that $\|x\|_X \leq C \|Tx\|_Y$ for all $x \in X$.

HINT: Prove that (1) $\implies$ (2) $\implies$ (3) $\implies$ (1).

Proof. (1) $\implies$ (2). If T is injective and range(T) is closed, then $\Gamma(T) = \{(x, Tx) : x \in X\}$ is closed in $X \times Y$. Therefore, $T : X \to \text{range}(T)$ is a bounded linear map. Since T is injective, this map is actually bijective from X to range(T). Therefore, T is a linaer homeomorphism.

(2) $\implies$ (3). Since T is a bijective bounded linear map, from X to range(T). There exists a constant $C \geq 0$ such that for each $y \in \text{range}(T)$ there exists a unique $x \in X$ such that $Tx = y$ and $\|x\| \leq C \|y\| = C \|Tx\|$. Since T is a bijection, $\|x\| \leq C \|Tx\|$ for all $x \in X$.

(3) $\implies$ (1). Let $x \in X$ be such that $Tx = 0$. It follows that $\|x\| \leq C \|Tx\| = 0$. Therefore, $x = 0$ and T is injective. To show that range(T) is closed, consider a convergent sequence

$\{y_n\} \subset \text{range}(T)$ with $y_n = Tx_n$. Since for any $n, m \in \mathbb{N}$ we have

$$\|x_n - x_m\| \leq C \|T(x_n - x_m)\| = C \|y_n - y_m\|,$$

$\{x_n\}$ is Cauchy. Since X is Banach, $x_n \to x$ for some $x \in X$. Therefore, for all $n \in \mathbb{N}$ we have

$$\|y_n - Tx\| = \|T(x_n - x)\| \leq \|T\| \|x_n - x\|,$$

and $y_n \to Tx$. Hence, range(T) is closed and the proof is complete.

□

Theorem. Let X and Y be Banach spaces over a common field. Then, the following subsets of $\mathcal{L}(X;Y)$ are open:

1. $\{T \in \mathcal{L}(X;Y) : T \text{ is surjective}\}$,
2. $\{T \in \mathcal{L}(X;Y) : T \text{ is injective with closed range}\}$,
3. $\mathcal{H}(X;Y) = \{T \in \mathcal{L}(X;Y) : T \text{ is a homeomorphism}\}$.

Proof. 1. Let $T \in \mathcal{L}(X;Y)$ be surjective. By open mapping theorem, there is $\delta > 0$ such that $B_Y(0, \delta) \subset TB_X(0, 1)$. By homogeneity we have $B_Y(0, r) \subset TB_X(0, \alpha r)$ for all $r > 0$ where $\alpha = \delta^{-1}$. Now let $S \in \mathcal{L}(X;Y)$ be such that $\|T - S\| < \beta < (2\alpha)^{-1}$. Claim S is surjective.

Let $y \in Y$, inductively construct sequences $\{x_n\}$ and $\{y_n\}$. First let $y_0 = y$. Then, $\|y_0\| \in B(0, 2\|y_0\|)$. Select $x_0 \in X$ be such that $Tx_0 = y_0$ and $\|x_0\| \leq 2\alpha \|y_0\|$. Suppose we have selected y_i, x_i for $0 \leq i \leq n$. Set $y_{n+1} = y_n - Sx_n$ and select x_{n+1} be such that $Tx_{n+1} = y_{n+1}$ and $\|x_{n+1}\| \leq 2\alpha \|y_{n+1}\|$. Then, we have

$$\|y_{n+1}\| = \|Tx_n - Sx_n\| \leq \|T - S\| \|x_n\| < 2\alpha\beta \|y_n\|$$

and

$$\|x_{n+1}\| = 2\alpha \|y_{n+1}\| \leq 2\alpha \|T - S\| \|x_n\| < 2\alpha\beta \|x_n\|.$$

Note that $2\alpha\beta < 1$ and X is Banach, define

$$x = \sum_{n=0}^{\infty} x_n = \lim_{N \to \infty} \sum_{n=0}^{N} x_n.$$

Also note that $\lim_{n \to \infty} y_n = 0$. It follows that

$$Sx = \sum_{n=0}^{\infty} Sx_n = \sum_{n=0}^{\infty} (y_n - y_{n+1}) = y_0 - \lim_{n \to \infty} y_{n+1} = y.$$

Therefore S is surjective and the set of surjective bounded linear maps are open.

2. Suppose $T \in \mathcal{L}(X;Y)$ is injective with closed range. Then, closed range theorem gives $C > 0$ such that $\|x\| \leq C\|Tx\|$ for all $x \in X$. Now suppose $S \in \mathcal{L}(X;Y)$ is such that $\|T - S\| < (2C)^{-1}$. Claim that S is also injective with closed range. Indeed,

$$\begin{aligned}\|x\| &\leq C\|Tx\| \leq C\|Sx\| + C\|(T-S)x\| \\ &\leq C\|Sx\| + \frac{1}{2}\|x\|.\end{aligned}$$

   This shows that $\|x\| \leq 2C\|Sx\|$ for all $x \in X$. By closed range theorem, S is injective with closed range. This implies that the set of injective bounded linear operator with closed range is open.

3. $\mathcal{H}(X;Y)$ is the intersection of surjective linear maps and injective maps with closed range. Therefore, $\mathcal{H}(X;Y)$ is open.

□

Theorem. Let X and Y be Banach spaces over a common field. Then the following holds.

1. The sets

$$\mathcal{L}_R(X;Y) = \{T \in \mathcal{L}(X;Y) : \text{there exists } S \in \mathcal{L}(Y;X) \text{ such that } ST = I_X\}$$

   and

$$\mathcal{L}_L(X;Y) = \{T \in \mathcal{L}(X;Y) : \text{there exists } S \in \mathcal{L}(Y;X) \text{ such that } TS = I_Y\}$$

   are open.

2. The following inclusion holds:

$$\mathcal{L}_L(X;Y) \subset \{T \in \mathcal{L}(X;Y) : T \text{ is surjective}\}$$

   and

$$\mathcal{L}_R(X;Y) \subset \{T \in \mathcal{L}(X;Y) : T \text{ is injective with closed range}\}.$$

3. The sets $\mathcal{L}_L(X;Y) \setminus \mathcal{L}_R(X;Y)$ and $\mathcal{L}_R(X;Y) \setminus \mathcal{L}_L(X;Y)$ are open.

Proof. 1. Let $T_0 \in \mathcal{L}_R$ and $S_0 \in \mathcal{L}(Y;X)$ be such that $T_0 S_0 = I_Y$. Note that $I_X \in \mathcal{H}(X)$ and when $\|P\| < 1$ for $P \in \mathcal{L}(X)$, we have $I_X + P \in \mathcal{H}(X)$. Suppose now $T \in \mathcal{L}(X;Y)$ and $\|T\| < \|S_0\|^{-1}$. It follows that $I_X + S_0 T \in \mathcal{H}(X)$. For such T, we then have

$$T_0 + T = T_0(I_X + S_0 T).$$

   Also,

$$(T_0 + T)(I_X + S_0 T)^{-1} S_0 = T_0(I_X + S_0 T)(I_X + S_0 T)^{-1} S_0 = T_0 S_0 = I_Y.$$

Therefore, $T_0 + T \in \mathcal{L}_R$ for $T \in B(T_0, \|S_0\|^{-1})$ and $\mathcal{L}_R$ is open.

Now let $T_0 \in \mathcal{L}_L$ and $S_0 \in \mathcal{L}(Y;X)$ be such that $S_0 T_0 = I_X$. Again, for $T \in \mathcal{L}(X;Y)$ with $\|T\| < \|S_0\|^{-1}$, we have

$$T_0 + T = (I_X + TS_0)T_0.$$

and

$$S_0(I_X + TS_0)^{-1}(T_0 + T) = I_X.$$

Therefore, $\mathcal{L}_R$ is also open.

2. Let $T \in \mathcal{L}_R$ and $S \in \mathcal{L}(Y;X)$ be such that $TS = I_Y$. Then for any $y \in Y$ let $x = Sy$. It follows that $Tx = TSy = y$. Also, $\|x\| \leq \|S\| \, \|y\|$ so the 4th item in open mapping theorem guarantees that T is surjective. Hence, $\mathcal{L}_L \subset \{T \in \mathcal{L}(X;Y) : T \text{ is surjective}\}$.

   Now let $T \in \mathcal{L}_L$ and $S \in \mathcal{L}(Y;X)$ such that $ST = I_X$. Now for any $x \in X$, we have $\|x\| = \|STx\| \leq \|S\| \, \|Tx\|$. Then the closed range theorem guarantees that T is injective with closed range. Hence, $\mathcal{L}_R \subset \{T \in \mathcal{L}_R(X;Y) : T \text{ is injective with closed range}\}$.

3. ***** TO-DO *****

□

1.3 Hahn-Banach theorem and duality

Theorem (Hahn-Banach theorem in $\mathbb{R}$)**.** Let X be a real vector space and suppose $p : X \to \mathbb{R}$ is such that

$$p(tx + (1-t)y) \leq tp(x) + (1-t)p(y)$$

for all $t \in [0,1]$ and $x, y \in X$.

Suppose Y subspace of X and $l : Y \to \mathbb{R}$ is a linear map such that $l \leq p$ on Y. Then there exists linear map $L : X \to \mathbb{R}$ such that $L \leq p$ on X and $L = l$ on Y.

Proof. Let

$$P = \{(Z, \lambda) : \; Y \subset Z \subset X, \, \lambda \text{ linear functional on } Z, \, \lambda \leq p \text{ on } Z \text{ and } l = \lambda \text{ on } Y\}$$

Define partial order $(Z_1, \lambda_1) \preceq (Z_2, \lambda_2)$ if and only if $Z_1 \subset Z_2$ and $\lambda_1 = \lambda_2$ on Z_1. It is easy to verify that this is a partial order. Towards using Zorn's Lemma, let $C \subset P$ be a chain and define

$$U = \bigcup_{(Z,\lambda) \in C} Z, \qquad \Lambda = \bigcup_{(Z,\lambda) \in C} \lambda.$$

It is easy to verify that (U, Λ) is an upper bound for the chain. By Zorn's Lemma, P has a maximal element (M, L). It remains to show thta $M = X$.

Suppose for contradiction that $M \neq X$. Pick $x_0 \in X \setminus M$. For any $x, y \in M$, we have

$$\begin{aligned}
\beta L(x) + \alpha L(y) &= L(\beta x + \alpha y) \\
&= \frac{1}{\alpha+\beta} L\left(\frac{\beta}{\alpha+\beta}x + \frac{\alpha}{\alpha+\beta}y\right) \\
&\leq (\alpha+\beta) p\left(\frac{\beta}{\alpha+\beta}x + \frac{\alpha}{\alpha+\beta}y\right) \\
&= (\alpha+\beta) p\left(\frac{\beta}{\alpha+\beta}(x - \alpha x_0) + \frac{\alpha}{\alpha+\beta}(y + \beta x_0)\right) \\
&\leq \beta p(x - \alpha x_0) + \alpha p(y + \beta x_0).
\end{aligned}$$

This implies that

$$\sup_{\substack{x \in M \\ \alpha > 0}} \frac{1}{\alpha}\left[L(x) - p(x - \alpha x_0)\right] \leq \inf_{\substack{y \in M \\ \beta > 0}} \frac{1}{\beta}\left[p(y + \beta x_0) - L(y)\right].$$

Note that $-p(-x_0) \leq \text{LHS}$ and $\text{RHS} \leq p(x_0)$, so $\text{LHS}, \text{RHS} < \infty$. Now pick $v \in \mathbb{R}$ such that $\text{LHS} \leq v \leq \text{RHS}$. For $x \in M$ and $0 < t \in \mathbb{R}$ we have

$$L(x) - tv \leq p(x - tv_0), \qquad L(x) + tv \leq p(x + tv_0).$$

Now define $\widehat{L} : M \oplus \mathbb{R}x_0 \to \mathbb{R}$ by $\widehat{L}(x + \alpha x_0) = L(x) + \alpha v$. It follows that $(M \oplus \mathbb{R}x_0, \widehat{L}) \in P$. However, $(M, L) \prec (M \oplus \mathbb{R}, \widehat{L})$, a contradiction. Therefore, $M = X$ and the proof is complete.

□

Theorem (Hahn-Banach theorem in $\mathbb{C}$)**.** Let X be complex vector space and suppose $p : X \to \mathbb{R}$ is such that

$$p(\alpha x + \beta y) \leq |\alpha|\, p(x) + |\beta|\, p(y)$$

for all $\alpha, \beta \in \mathbb{C}$ such that $|\alpha| + |\beta| = 1$ and $x, y \in X$.

Suppose Y subspace of X and $l : Y \to \mathbb{C}$ is a linear map such that $|l| \leq p$ on Y. Then there exists linear map $L : X \to \mathbb{C}$ such that $|L| \leq p$ on X and $L = l$ on Y.

Proof. Define $\lambda : Y \to \mathbb{R}$ by $\lambda(x) = \operatorname{Re}(l(x))$. Note that

$$\lambda(ix) = \operatorname{Re}(il(x)) = -\operatorname{Im}(l(x)).$$

This implies that $l(x) = \lambda(x) - i\lambda(ix)$. Now treat X and Y as vector space over $\mathbb{R}$ and apply Hahn-Banach theorem in $\mathbb{R}$ to extend λ to $\Lambda : X \to \mathbb{R}$ that agrees with λ on Y.

Define $L : X \to \mathbb{C}$ by $L(x) = \Lambda(x) - i\Lambda(ix)$. It remains to show that $|L| \leq p$. For $x \in X$, write $L(x) = |L(x)| e^{i\theta}$ for some $\theta \in \mathbb{R}$. It follows that

$$\begin{aligned} |L(x)| &= L(x)e^{-i\theta} \\ &= \Lambda(e^{-i\theta}) - i\Lambda(ie^{-i\theta}x) \\ &= \Lambda(e^{-i\theta}x) \\ &\leq p(e^{-i\theta x}) \\ &\leq \left|e^{-i\theta}\right| p(x) \\ &= p(x), \end{aligned}$$

as desired.

□

Theorem (Hahn-Banach theorem for bounded linear functionals)**.** Let X be a normed vector space over $\mathbb{F}$ and Y a subspace of X. If $\lambda \in Y^*$ then there exists $\Lambda \in X^*$ such that $\Lambda = \lambda$ on Y and the operator norm $\|\lambda\|_{Y^*} = \|\Lambda\|_{X^*}$.

Proof. Consider $p : X \to \mathbb{R}$ where $p(x) = \|\lambda\|_{Y^*} \|x\|$. Apply Hahn-Banach theorem.

□

Next we show some useful implications of Hahn-Banach theorem.

Theorem. Let X be a normed vector space and fix $x \in X$. Then the following holds:

1. There exists $\lambda \in X^*$ such that $\|\lambda\| = \|x\|$ and
$$\lambda(x) = \|\lambda\| \|x\| = \|x\|^2 .$$
2. We have
$$\|x\| = \max_{\substack{w \in X^* \\ \|w\|=1}} |w(x)| .$$
3. $x = 0$ if and only if $w(x) = 0$ for all $w \in X^*$.

Proof.
1. Let $Y = \mathbb{F}x$ and define $\lambda \in Y^*$ by $\lambda(ax) = a \|x\|^2$. Apply Hahn-Banach theorem.
2. Suppose $x \neq 0$. Define $w = \frac{\lambda}{\|x\|}$ then it follows that $|w(x)| = \|x\|$.
3. Follows directly from (2).

□

Proposition. Let X be normed vector space. Then the mapping $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{F}$ by $(w, x) \mapsto w(x)$ is a bilinear map. That is, $\langle \cdot, \cdot \rangle \in \mathcal{L}(X^*, X; \mathbb{F})$. Moreover, if $X \neq \{0\}$, then $\|\langle \cdot, \cdot \rangle\| = 1$.

Proof. It is easy to see that $\langle \cdot, \cdot \rangle$ is bilinear. For boundedness,

$$|\langle w, x \rangle| = |w(x)| \leq \|w\| \, \|x\| \, .$$

Hence, $\|\langle \cdot, \cdot \rangle\| \leq 1$. Meanwhile, pick some $x \in X$ with $\|x\| = 1$. It follows that

$$1 = \|x\| = \max_{\substack{w \in X^* \\ \|w\|=1}} |w(x)| \leq \|\langle \cdot, \cdot \rangle\| \, .$$

Therefore, $\|\langle \cdot, \cdot \rangle\| = 1$. □

Definition (Norming set)**.** Let X be normed vector space and $E \subset X$, $W \subset X^*$. Say W is a **norming set** for E if

$$\|x\| = \sup_{\substack{w \in W \\ \|w\|=1}} |\langle w, x \rangle|$$

for all $x \in E$.

Proposition. Let X be normed vector space and $S \subset X$ be a separable set. Let W be a norming set for S. Then, there exists $\{w_n\}_{n=0}^{\infty} \subset W$ such that $\|w_n\| = 1$, and the sequence is norming for S. That is,

$$\|x\| = \sup_{n \in \mathbb{N}} |\langle w_n, x \rangle| \, .$$

Proof. Let $\{v_n\}_{n=0}^{\infty} \subset S$ be dense. For any $n, k \in \mathbb{N}$, choose $w_{n,k} \in W$ with $\|w_{n,k}\| = 1$ such that

$$(1 - 2^{-k}) \|v_n\| \leq |w_{n,k}, v_n| \, .$$

Let $x \in S$ and $0 < \varepsilon < 1$ be arbitrary. Pick $v_n \in S$ such that $\|v_n - x\| < \varepsilon$ and pick $j \in \mathbb{N}$ such that $2^{-j} < \varepsilon$. Then,

$$\begin{aligned}
(1 - \varepsilon) \|x\| &\leq (1 - 2^{-j}) \|x\| \\
&\leq (1 - 2^{-j}) \|v_n\| + (1 - 2^{-j}) \|v_n - x\| \\
&\leq |\langle w_{n,j}, v_j \rangle| + \varepsilon \\
&\leq |\langle w_{n,j}, x \rangle| + |\langle w_{n,j}, x - v_n \rangle| + \varepsilon \\
&\leq |\langle w_{n,j}, x \rangle| + 2\varepsilon.
\end{aligned}$$

This shows that $\{w_{n,k}\}_{n,k=0}^{\infty}$ is a norming sequence.

□

Theorem. Let X be normed vector space and define $J : X \to X^{**}$ by $\langle Jx, w \rangle = \langle w, x \rangle = w(x)$. Then the following holds:

1. $J \in \mathcal{L}(X, X^{**})$.
2. J is an isometric embedding. In particular, it is injective.
3. $\operatorname{range}(J) \subset X^{**}$ is a norming set for X^*.
4. X is Banach if and only if $\operatorname{range}(J)$ is closed.

Proof. Note that we have

$$\begin{aligned} \|Jx\|_{X^{**}} &= \sup\{|\langle Jx, w\rangle| : w \in X^* \text{ and } \|w\| \le 1\} \\ &= \sup\{|\langle w, x\rangle| : w \in X^* \text{ and } \|w\| \le 1\} \\ &= \|x\|, \end{aligned}$$

where the last step is by a previous theorem that shows the existence of $w \in X^*$ such that $\|w\| = 1$ and $|w(x)| = \|x\|$. This implies (1) and (2). Now we know X is isometrically isomophic to $\operatorname{range}(J) \subset X^{**}$. Therefore, X is Banach if and only if $\operatorname{range}(J)$ is Banach. However, $X^{**} = \mathcal{L}(X^*, \mathbb{F})$ is Banach, so $\operatorname{range}(J)$ is Banach if and only if $\operatorname{range}(J)$ is closed. This implies (4).

To show (3), note that we have

$$\begin{aligned} \|w\|_{X^*} &= \sup\{|\langle w, x\rangle| : x \in X \text{ and } \|x\| \le 1\} \\ &= \sup\{|\langle Jx, w\rangle| : x \in X \text{ and } \|x\| \le 1\} \\ &= \sup\{|\langle v, w\rangle| : v \in \operatorname{range}(J) \text{ and } \|v\|_{X^{**}} \le 1\}. \end{aligned}$$

This shows (3), completing the proof. □

2 Differential calculus

2.1 Inverse and implicit function theorem

2.1.1 Local injectivity and surjectivity theorem

Theorem (Local injectivity theorem)**.** Let X and Y be Banach spaces, $z \in U \subset X$ with U open. Let $f : U \to Y$ differentiable with Df continuous at z. Suppose $Df(z) \in \mathcal{L}(X; Y)$ injective with closed range. Then for any $0 < \varepsilon < 1$, there exists $r > 0$ such that

1. $B[z, r] \subset U$.
2. $Df(x)$ injective with closed range for all $x \in B[z, r]$.
3. If $x, y \in B(z, r)$, then
$$(1-\varepsilon)\left\|Df(z)(x-y)\right\| \le \left\|f(x)-f(y)\right\| \le (1+\varepsilon)\left\|Df(z)(x-y)\right\|.$$
4. The restriction $f : B(z, r) \to f(B(z, r))$ is bi-Lipschitz homeomorphism.

Proof. Since $Df(z)$ injective with closed range, there exists $\theta > 0$ such that
$$\theta \left\|h\right\| \le \left\|Df(z)h\right\|$$
for all $h \in X$. Since the set of bounded linear operator that is injective with closed range is open, there exists $\delta > 0$ such that $\left\|Df(z) - T\right\| < \delta$ implies T is injective with closed range.

Now let $0 < \varepsilon < 1$. Note that Df is continuous at z, so we can select $r > 0$ so small that $B[z, r] \subset U$, and $x \in B[z, r]$ implies
$$\left\|Df(x) - Df(z)\right\| < \min\left\{\delta, \theta\varepsilon\right\}.$$
In particular, $Df(x)$ is injective with closed range for all $x \in B[z, r]$. By the mean value theorem, for any $x, y \in B(x, r)$
$$\begin{aligned}\left\|f(x)-f(y)-Df(z)(x-y)\right\| &\le \sup_{w\in B(z,r)} \left\|Df(w)-Df(z)\right\|\left\|x-y\right\| \\ &\le \theta\varepsilon\left\|x-y\right\| \\ &\le \varepsilon\left\|Df(z)(x-y)\right\|.\end{aligned}$$
It follows that
$$(1-\varepsilon)\left\|Df(z)(x-y)\right\| \le \left\|f(x)-f(y)\right\| \le (1+\varepsilon)\left\|Df(z)(x-y)\right\|,$$
as desired.

This also implies that
$$(1-\varepsilon)\theta\left\|x-y\right\| \le \left\|f(x)-f(y)\right\| \le (1+\varepsilon)\left\|Df(z)\right\|\left\|x-y\right\|,$$
so the restriction of f on $B(z, r)$ is a bi-Lipschitz homeomorphism.

□

Theorem (Local surjectivity theorem)**.** Let X and Y be Banach spaces, $z \in U \subset X$ with U open. Let $f : U \to Y$ differentiable with Df continuous at z. Suppose $Df(z) \in \mathcal{L}(X;Y)$ surjective. Then there exists $r_0,\ \gamma > 0$ such that

1. $B_X[z, r_0] \subset U$.
2. $Df(x)$ surjective for all $x \in B_X[z, r_0]$.
3. $B_Y[f(z), \gamma r] \subset f(B_X[z, r])$ for all $0 \leq r \leq r_0$.

Proof. **Step 1: reduction**

WLOG we can suppose $z = 0$ and $f(z) = 0$. Indeed, suppose the theorem holds in this special case and consider the general hypotheses. Define $g : U - z \to Y$ by $g(x) = f(x + z) - f(z)$. Then note that $z \in U$ so $0 \in U - z$. Also, $g(0) = 0$ and $Dg(x) = Df(x + z)$. Therefore, $Dg(0) = Df(z)$ is surjective. Now apply the special case to g to obtain $r_0,\ \gamma > 0$ such that

1. $B_X[0, r_0] \subset U - z$.
2. $Dg(x)$ is surjective for $x \in B_X[0, r_0]$.
3. $B_Y[0, \gamma r] \subset g(B_X[0, r])$ for all $0 \leq r \leq r_0$.

It follows that

1. $B_X[z, r_0] \subset U$.
2. $Df(x)$ surjective for all $x \in B_X[z, r_0]$.
3. $B_Y[f(z), \gamma r] \subset f(B_X[z, r])$ for all $0 \leq r \leq r_0$.

Step 2: set up

Suppose $z = 0$ and $f(z) = 0$. Write $T = Df(0) \in \mathcal{L}(X;Y)$, which is surjective. Then by open mapping theorem, there exists $M > 0$ such that for each $y \in Y$, there is $x \in X$ such that $Tx = y$ and $\|x\| \leq M \|y\|$. We also know the set of surjective linear maps in $\mathcal{L}(X;Y)$ is open, so there exists $\varepsilon > 0$ such that $\|S - T\| < \varepsilon$ implies S surjective.

Let $\gamma = \frac{1}{2M} > 0$ and use continuity to pick $r_0 > 0$ such that $B_X[0, r_0] \subset U$ and

$$\|Df(x) - T\| < \min\{\gamma, \varepsilon\}$$

for all $x \in B_X[0, r_0]$. In particular, $Df(x)$ is surjective for all $x \in B_X[0, r_0]$.

Finally, by mean value theorem, if $u, v \in B_X[0, r_0]$, then

$$\|f(u) - f(v) - T(u, v)\| \leq \|u - v\| \sup_{w \in B_X[0, r_0]} \|Df(w) - T\| \leq \gamma \|u - v\|. \tag{$*$}$$

□

Step 3: completion

Fix $0 < r \leq r_0$ and $y \in B_Y[0, \gamma r]$, we want to show $y \in f(B_X[0,r])$. We now inductively construct a sequence $\{x_n\}_{n=0}^{\infty}$. Let $x_0 = 0$.

(Base case) Pick $x_1 \in X$ such that $Tx_1 = y$ and $\|x_1\| \leq M\|y\| \leq M\gamma r = \frac{r}{2} < r$. Hence $x_1 \in B(0,r)$. Moreover, we have

$$Tx_1 = y = y - f(x_0) + Tx_0,$$

and

$$\|x_1 - x_0\| = \|x_1\| \leq M\|y\|.$$

(Inductive step) Now suppose $1 \leq n \in \mathbb{N}$ and we constructed $\{x_m\}_{m=0}^{n} \subset B_X(0,r)$ such that $Tx_m = y - f(x_{m-1}) + Tx_{m-1}$, and $\|x_m - x_{m-1}\| \leq 2^{-m+1}M\|y\|$ for all $1 \leq m \leq n$. Let $h \in X$ such that $Th = y - f(x_n)$ and $\|h\| \leq M\|y - f(x_n)\|$. Set $x_{n+1} = x_n + h$. Note that now

$$Tx_{n+1} = Tx_n + Th = y - f(x_n) + Tx_n,$$

and

$$\|x_{n+1} - x_n\| = \|h\| \leq M\|y - f(x_n)\|.$$

Now we have $y = f(x_{n-1}) + T(x_n - x_{n-1})$, so

$$y - f(x_n) = f(x_{n-1}) - f(x_n) - T(x_{n-1} - x_n).$$

By $(*)$, we have $\|y - f(x_n)\| \leq \gamma\|x_{n-1} - x_n\|$, and thus

$$\|x_{n+1} - x_n\| \leq M\|y - f(x_n)\| \leq M\gamma\|x_{n-1} - x_n\| \leq 2^{-n}M\|y\|.$$

In turn,

$$\|x_{n+1}\| = \|x_{n+1} - x_0\| \leq \sum_{k=0}^{n}\|x_{k+1} - x_k\| < M\|y\|\sum_{k=0}^{\infty}2^{-k} = 2M\|y\|.$$

However, $\gamma = \frac{1}{2M}$. This then implies that $2M\|y\| = \frac{1}{\gamma}\|y\| \leq r$ since $y \in B_Y[0,\gamma r]$. Therefore, $x_{n+1} \in B_X(0,r)$, completing the inductive construction.

Note that $\{x_n\}_{n=0}^{\infty}$ is Cauchy and hence $x_n \to x$ for some $x \in B_X[0,r]$ as $n \to \infty$. This then implies that

$$\begin{aligned} Tx = \lim_{n\to\infty} Tx_n &= \lim_{n\to\infty}[y - f(x_{n-1}) - Tx_{n-1}] \\ &= y - f(x) + Tx. \end{aligned}$$

It follows that $y = f(x)$. Since $y \in B_Y[0,\gamma r]$ is arbitrary, we have that $B_Y[0,\gamma r] \subset f(B_X[0,r])$ for all $0 \leq r \leq r_0$, completing the proof for the theorem.

Remark. The argument in step 3 is in fact **Newton's method**. The idea is as follows: Suppose $f : X \to Y$ is C^1 and $Df(x)$ is an isomophism for all x. Say we want to solve the equation $y = f(x)$ for x and we know $y_0 = f(x_0)$ and $y \approx y_0$. Now we have

$$y = f(x_0) + Df(x_0)(x - x_0) = y_0 + Df(x_0)(x - x_0).$$

It follows that $y - y_0 = Df(x_0)(x - x_0)$ and thus $x = x_0 + [Df(x_0)]^{-1}(y - y_0)$.

This gives the following **algorithm** to solve an equation. Define

$$x_{n+1} = x_n + [Df(x_n)]^{-1}(y - f(x_n)).$$

Suppose $x_n \to x$ as $n \to \infty$, then

$$x = x + [Df(x)]^{-1}(y - f(x))$$

and thus $y = f(x)$.

Definition (diffeomorphism)**.** Let X and Y be normed vector spaces and suppose that $\emptyset \neq U \subset X$ is open. Let $f : U \to Y$. For $k \geq 1$, say f is a C^k diffeomorphism if

1. $f : U \to f(U)$ homeomorphism with $f(U) \subset Y$ open.
2. $f \in C^k(U; Y)$.
3. $f^{-1} \in C^k(f(U); X)$.

If f is a C^k diffeomorphism for all $k \geq 1$, say f is a smooth diffeomorphism.

Theorem. Let X and Y be Banach spaces and $\emptyset \neq U \subset X$ open. Suppose $f \in C^k(U; Y)$ for some $k \geq 1$ and f is a homeomorphism from U to $f(U)$. Then the following are equivalent:

1. f is a C^k diffeomorphism.
2. $Df(x) \in \mathcal{L}(X; Y)$ is an isomorphism for all $x \in U$.

Proof. ***** TO-DO ***** □

2.1.2 Inverse function theorem

Theorem (Inverse function theorem)**.** Let X and Y be Banach spaces, $U \subset X$ open and $x_0 \in U$. Suppose $f : U \to Y$ differentiable, Df continuous at x_0, $Df(x_0)$ linear homeomorphism. Then there exists bounded and open $V \subset U$ with $x_0 \in V$ such that

1. $f : V \to f(V)$ is bi-Lipschitz homeomorphism, $Df(x)$ linear homeomorphism for all $x \in V$, $f(V) \subset Y$ bounded and open, $f^{-1} : f(V) \to V$ differentiable with

$$Df^{-1}(y) = [Df(f^{-1}(y))]^{-1}$$

for all $y \in f(V)$ and Df^{-1} is continuous at $f(x_0)$. Also, there exists C_0, $C_1 > 0$ such that

$$C_0 \leq \|Df(x)\| \leq C_1$$

for all $x \in V$, and
$$\frac{1}{C_1} \leq \left\|Df^{-1}(y)\right\| \leq \frac{1}{C_0}$$
for all $y \in f(V)$.

2. If $f \in C^k(U;Y)$ for some $1 \leq k \leq \infty$, then $f^{-1} \in C^k(f(V);X)$. In particular, f is a local C^k diffeomorphism at x_0.
3. If $f \in C^k(U;Y)$ for $1 \leq k \in \mathbb{N}$, then there exists open $V_k \subset V$ such that $x_0 \in V_k$, $f \in C_b^k(V_k;Y)$ and $f^{-1} \in C_b^k(f(V_k);X)$.

Proof. ***** TO-DO ***** □

2.1.3 Implicit function theorem

Theorem (Implicit function theorem)**.** Let X and Y be Banach spaces, $U \subset X \times Y$ be open with $(x_0, y_0) \in U$, and suppose $f : U \to Z$ is differentiable in U with Df continuous at (x_0, y_0). Further suppose $z_0 = f(x_0, y_0)$ and $D_2 f(x_0, y_0) \in \mathcal{L}(Y;Z)$ is an isomorphism. Then there exists open sets $x_0 \in V \subset X$, $z_0 \in W \subset Z$, $y_0 \in S \subset Y$, and $g \in C_b^{0,1}(V \times W;Y)$ such that the following holds:

1. $g(x_0, z_0) = y_0$ and $(x, g(x,z)) \in V \times S \subset U$ for all $(x,z) \in V \times W$. Also, g is differentiable on $V \times W$ and Dg continuous at (x_0, z_0).
2. $f(x, g(x,z)) = z$ for all $(x,z) \in V \times W$. Moreover, if $(x,y) \in V \times S$ such that $f(x,y) = z$ for some $z \in W$, then $y = g(x,z)$.
3. $D_2 f(x, g(x,z))$ is an isomorphism for all $(x,z) \in V \times W$, and
$$\begin{aligned} D_1 g(x,z) &= -\left[D_2 f(x, g(x,z))\right]^{-1} D_1 f(x, g(x,z)), \\ D_2 g(x,z) &= \left[D_2 f(x, g(x,z))\right]^{-1}. \end{aligned}$$
4. If $f \in C^k$ then $g \in C^k$ too for $1 \leq k \leq \infty$. If k finite and $f \in C_b^k$ then the sets can be picked such that $g \in C_b^k$.

Proof. ***** TO-DO ***** □

2.1.4 Left and right inverse function theorems and flattening maps

Theorem (left inverse function theorem)**.** Let X and Y be Banach spaces. Suppose that $\emptyset \neq U \subset X$ is open and $f \in C^k(U;Y)$ for some $1 \leq k \leq \infty$. Let $x_0 \in U$ and $y_0 = f(x_0) \in Y$, and suppose that $\{0\} \neq \operatorname{range} Df(x_0) \subset Y$. Then the following are equivalent:

1. The map $Df(x_0) \in \mathcal{L}(X;Y)$ is injective with $\operatorname{range} Df(x_0)$ closed and complemented in Y. That is, $Df(x_0) \in \mathcal{L}_L(X;Y)$.

2. There exist nontrivial closed subspace $Y_0, Y_1 \subset Y$ with $Y = Y_0 \oplus Y_1$ and open sets $x_0 \in \widetilde{U} \subset U$ and $0 \in S \subset Y_1$ such that the map $G : \widetilde{U} \times S \to Y$ given by $G(x, y) = f(x) + y$ is a C^k diffeomorphism onto its image. Moreover, we have that $DG(x_0, 0)(X \times \{0\}) = Y_0$ and the restriction $DG(x_0, 0)|_{X\times\{0\}} : X \times \{0\} \to Y_0$ is an isomorphism.

3. There exist nontrivial closed subspace $Y_0, Y_1 \subset Y$ with $Y = Y_0 \oplus Y_1$, open sets $x_0 \in \widetilde{U} \subset U$, $0 \in S \subset Y_1$, and $W \subset Y$. and a map $F \in C^k(W; X \times Y_1)$ such that F is a C^k diffeomorphism from W to $F(W) = \widetilde{U} \times S$, $f(\widetilde{U}) \subset W$, and
$$F(f(x)) = (x, 0)$$
for all $x \in \widetilde{U}$. Moreover, $DF(y_0)(Y_0) = X \times \{0\}$ and the restriction $DF(y_0)|_{Y_0} : Y_0 \to X \times \{0\}$ is an isomorphism.

4. There exist open sets $x_0 \in \widetilde{U} \subset U$ and $W \subset Y$ with $f(\widetilde{U}) \subset W$, and a map $g \in C^k(W; X)$ such that $g(f(x)) = x$ for all $x \in \widetilde{U}$.

5. There exists an open set $x_0 \in \widetilde{U} \subset U$ such that $Df(x) \in \mathcal{L}(X; Y)$ is injective with range $Df(x)$ closed and complemented in Y for each $x \in \widetilde{U}$.

Moreover, in any and hence every case, the Y_i spaces from the second and third item can be taken with $Y_0 = \operatorname{range} Df(x_0)$, and the maps g, F, G can be taken to be related via $F^{-1} = G$ and $g = \pi_X \circ F$.

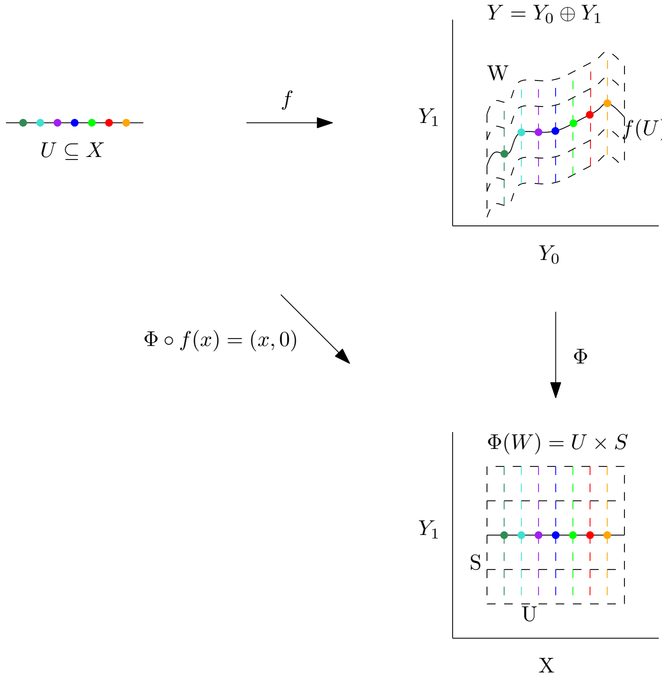
Proof. (1) $\implies$ (2). Write $Y_0 = \operatorname{range} Df(x_0)$. Since $Df(x_0)$ is injective with range $Df(x_0)$ closed and complemented in Y, there exists nontrivial subspace Y_1 such that $Y_0 \oplus Y_1 = Y$.

Now claim $DG(x_0, 0) \in \mathcal{L}(X \times Y_1; Y)$ is an isomorphism. Indeed, we have
$$DG(x_0, 0)(h, k) = Df(x_0)h + k.$$
It follows that the range $DG(x_0, 0) = \operatorname{range} Df(x_0) + Y_1 = Y$, so $DG(x_0, 0)$ is surjective. On the other hand, suppose $DG(x_0, 0)(h, k) = 0$. The direct sum composition implies that $k = 0$ and $Df(x_0)h = 0$. However, $Df(x_0)$ is injective so we also have $h = 0$. This implies that $DG(x_0, 0)$ is injective, and $DG(x_0, 0) \in \mathcal{L}(X \times Y_1; Y)$ is indeed an isomophism. It is also clear that $DG(x_0, 0)(X \times \{0\}) = \operatorname{range} Df(x_0) = Y_0$, and $DG(x_0, 0)|_{X\times\{0\}}$ is an isomophism.

Now the inverse function theorem then applies and gives a bounded and open set $V \subset U \times Y_1$ with $(x_0, 0) \in V$ and G is a C^k diffeomorphism from V to $G(V)$. We can then choose open sets $x_0 \in \widetilde{U} \subset U$ and $0 \in S \subset Y_1$ so that $\widetilde{U} \times S \subset V$. It follows that $G : \widetilde{U} \times S \to Y$ is a C^k diffeomorphism onto its image.

(2) $\implies$ (3). Define $W = G(\widetilde{U} \times S) \subset Y$ and let $F = G^{-1}$. It follows that F is a C^k diffeomorphism from W to $F(W) = \widetilde{U} \times S$. We also know that $f(\widetilde{U}) \subset G(\widetilde{U} \times S) = W$. Since $G(x, 0) = f(x)$ for all $x \in \widetilde{U}$, we have
$$F(f(x)) = (x, 0)$$

Depiction of the Left Inverse Function Theorem



for all $x \in \widetilde{U}$.

Moreover, note that $DF(y_0) = [DF(x_0, 0)]^{-1}$. Therefore, $DF(y_0)(Y_0) = X \times \{0\}$ and the restriction $DF(y_0)|_{Y_0} : Y_0 \to X \times \{0\}$ is an isomophism.

(3) $\implies$ (4). Let $g = \pi \circ F$ where π is the projection onto the first coordinate. Then $g \in C^k(W; X)$ and $g(f(x)) = x$ for all $x \in \widetilde{U}$

(4) $\implies$ (5). Since $g(f(x)) = x$ for all $x \in \widetilde{U}$, we have

$$D(g \circ f)(x) = Dg(f(x))Df(x) = I_X.$$

It follows that $Df(x) \in \mathcal{L}_L(X; Y)$ for all $x \in \widetilde{U}$, and thus $Df(x) \in \mathcal{L}(X; Y)$ is injective with range $Df(x)$ closed and complemented in Y for all $x \in \widetilde{U}$.

(5) $\implies$ (1). Trivially true since $x_0 \in \widetilde{U}$. □

Theorem (flattening map)**.** Let X and Y be Banach spaces, and suppose that $Y_0, Y_1 \subset Y$ are nontrivial clsoed subspaces such that $Y = Y_0 \oplus Y_1$. Suppose further that X and Y_0 are isomophic. Let $M \subset Y$ and $y_0 \in M$, and let $1 \leq k \leq \infty$. Then the following are equivalent:

1. There exist open sets $V \subset X$ and $U \subset Y$ with $y_0 \in U$ and $f \in C^k(U; Y)$ such that $f : V \to f(V) = U \cap M$ is a homeomorphism and at $x_0 = f^{-1}(y_0) \in V$ we have that $Df(x_0)$ is injective with range $Df(x_0) = Y_0$.
2. There exists an open set $W \subset Y$ with $y_0 \in W$ and $F \in C^k(W; X \times Y_1)$ such that F is a C^k diffeomorphism from W to $F(W)$, $DF(y_0)(Y_0) = X_0 \times \{0\}$, and
$$F(W \cap M) = F(W) \cap [X \times \{0\}].$$
3. There exists an open set $W \subset Y$ with $y_0 \in W$ and $\Phi \in C^k(W; Y)$ such that Φ is a C^k diffeomorphism from W to $\Phi(W)$, $D\Phi(y_0)(Y_0) = Y_0$, and
$$\Phi(W \cap M) = \Phi(W) \cap Y_0.$$

Moreover, if either the second or the third items hold, then the set V from the first item can be chosen such that $Df(x)$ is injective with range $Df(x)$ closed and complemented in Y for each $x \in V$.

Proof. (1) $\implies$ (2). By the left inverse function theorem, there exists open sets $x_0 \in \widetilde{V} \subset V$ and $0 \in S \subset Y_1$ such that the map $G : \widetilde{V} \times S \to Y$ given by $G(x, y) = f(x) + y$ is a C^k diffeomorphism onto its image $G(\widetilde{V} \times S) \subset Y$.

Note that f is a homeomorphism from V to $f(V) = U \cap M$ and $\widetilde{V}$ is open, it follows that there exists open set $R \subset U$ such that
$$G(\widetilde{V} \times \{0\}) = f(\widetilde{V}) = M \cap R.$$

Now define $W = G(\widetilde{V} \times S) \cap R \subset U$. We then have $W \subset R$ and $G^{-1}(M \cap R) = \widetilde{V} \times \{0\}$. Also let F be the inverse of G restricted to W, then F is a C^k diffeomorphism from W to $F(W)$ such that $DF(y_0)(Y_0) = X_0 \times \{0\}$. Using the fact that G is a diffeomorphism, it is also easy to verify that
$$F(W \cap M) = F(W) \cap [X \times \{0\}].$$

(2) $\implies$ (1). Let $U = W$ and $V = \{x \in X : (x, 0) \in F(W)\}$. Also define $f : V \to Y$ by
$$f(x) = F^{-1}(x, 0).$$

It is clear that $f \in C^k(V; Y)$, $f(V) = W \cap M$, and f is a homeomorphism. Since $y_0 \in W \cap M$, there exists a unique $x_0 \in X$ such that $f(x_0) = y_0$. We can then compute
$$Df(x_0)h = DF^{-1}(x_0, 0)(h, 0) = [DF(y_0)]^{-1}(h, 0).$$

It follows that
$$\operatorname{range} Df(x_0) = \operatorname{range}[DF(y_0)]^{-1}(X \times \{0\}) = Y_0.$$

Finally, set $H : W \to V$ by $H = \pi \circ F$, where π is the projection onto the first coordinate. Then $H(f(x)) = x$ for all $x \in V$, and hence

$$DH(f(x))Df(x) = I_X$$

for all $x \in V$. This implies that $Df(x) \in \mathcal{L}_L$, and thus $Df(x)$ is injective with range $Df(x)$ closed and complemented in Y for all $x \in V$, completing the proof for (2) $\implies$ (1).

(2) $\implies$ (3). This is clear since X is isomophic to Y_0. □

Theorem (right inverse function theorem)**.** Let X and Y be Banach spaces. Suppose that $\emptyset \neq U \subset X$ is open and $f \in C^k(U; Y)$ for some $1 \leq k \leq \infty$. Let $x_0 \in U$, $y_0 = f(x_0) \in Y$, and suppose that $\{0\} \neq \ker Df(x_0) \subset X$. Then the following are equivalent:

1. The map $Df(x_0) \in \mathcal{L}(X; Y)$ is surjective with $\ker Df(x_0)$ complemented in X. That is, $Df(x_0) \in \mathcal{L}_R(X; Y)$.
2. There exist nontrivial closed subspace $X_0, X_1 \subset X$ with $X = X_0 \oplus X_1$ and $P \in \mathcal{L}(X)$ the projection onto X_0 and open sets $x_0 \in \widetilde{U} \subset U$, $y_0 \in W \subset Y$, and $Px_0 \in V \subset X_0$ such that the map $G : \widetilde{U} \to Y \times X_0$ given by $G(x) = (f(x), Px)$ is a C^k diffeomorphism onto $W \times V \subset Y \times X_0$. Moreover, for $x \in \widetilde{U}$, we have that $DG(x)(X_1) = Y \times \{0\}$, and the restriction $DG(x)|_{X_1} : X_1 \to Y \times \{0\}$ is an isomorphism.
3. There exist nontrivial closed subspace $X_0, X_1 \subset X$ with $X = X_0 \oplus X_1$ and $P \in \mathcal{L}(X)$ the projection onto X_0, open sets $x_0 \in \widetilde{U} \subset U$, $y_0 \in W \subset Y$, and $Px_0 \in V \subset X_0$, and a map $F \in C^k(W \times V; X)$ such that F is a C^k diffeomorphism from $W \times V$ to $F(W \times V) = \widetilde{U}$, $F(y_0, Px_0) = x_0$, and
$$f(F(y, v)) = y$$
for all $(y, v) \in W \times V$. Moreover, for all $(w, v) \in W \times V$ we have that $DF(w, v)(Y \times \{0\}) = X_1$ and the restriction $DF(w, v)|_{Y \times \{0\}} : Y \times \{0\} \to X_1$ is an isomorphism.
4. There exists an open set $y_0 \in W \subset Y$ and a map $g \in C^k(W; X)$ such that $g(y_0) = x_0$, $g(W) \subset U$, and $f(g(y)) = y$ for all $y \in W$.
5. There exists an open set $x_0 \in \widetilde{U} \subset U$ such that $Df(x) \in \mathcal{L}(X; Y)$ is surjective with $\ker Df(x)$ complemented in X for each $x \in \widetilde{U}$.

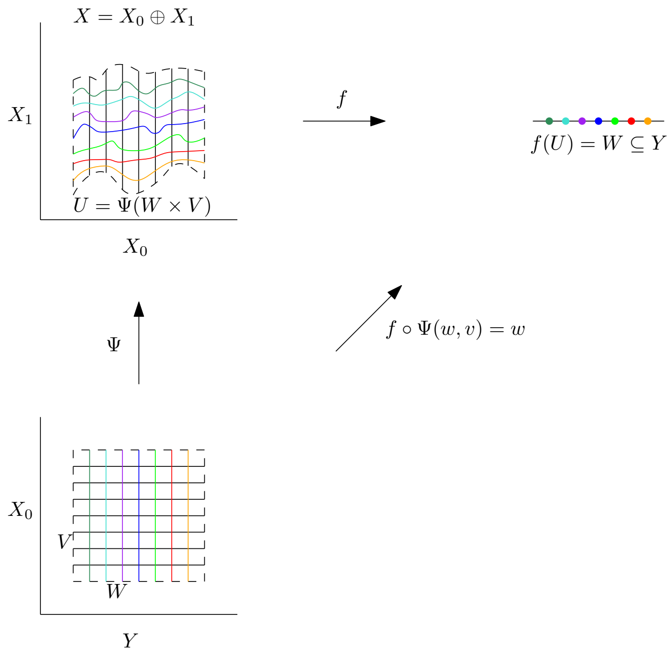
Moreover, in any and hence every case, the X_i spaces from the second and third items can be taken with $X_0 = \ker Df(x_0)$, and the maps g, F, G can be aken to be related via $F^{-1} = G$ and $g(y) = F(y, Px_0)$.

Proof. (1) $\implies$ (2). Let $X_0 = \ker Df(x_0)$ and P the projection on to X_0, then there exists nontrivial closed subspace $X_1 \subset X$ such that $X = X_0 \oplus X_1$.

Now claim $DG(x_0) \in \mathcal{L}(X; Y \times X_0)$ is an isomophism. Indeed, we have

$$DG(x_0)h = (Df(x_0)h, Ph).$$

Depiction of the Right Inverse Function Theorem



For injectivity, consider $D\Phi(x_0)h = (0, 0)$. Then, $Df(x_0)h = 0$ and $Ph = 0$. It follows that $h \in \ker Df(x_0) = X_0$. Since P restricted on range $P = X_0$ is the identity, we must have $h = 0$. Hence, $D\Phi(x_0)$ is injective. For surjectivity, consider $(y, k) \in Y \times X_0$. Since $Df(x_0)$ is surjective with $\ker Df(x_0)$ complemented in X, there exists $h \in X$ such that $Ph = 0$ and $Df(x_0)h = y$. Then,

$$D\Phi(x_0)(h + k) = (D\Phi(x_0)(h + k), P(h + k)) = (D\Phi(x_0)h, Pk) = (y, k),$$

where we utilized the fact that $k \in X_0 = \ker Df(x_0)$. This implies that $D\Phi(x_0)$ is a linear homeomorphism. It then follows from the inverse function theorem that there exists open set $y_0 \in W \subset Y$, $Px_0 \in V \subset X$, and $x_0 \in \widetilde{U} \subset U$ such that $G : \widetilde{U} \to Y \times X_0$ is a C^k diffeomorphism onto $G(\widetilde{U}) = W \times V$.

Note that $Df(x_0)|_{X_1} \in \mathcal{L}(X_1; Y)$ is an isomophism. Hence, shrinking $\widetilde{U}$ if necessary, we can assume $Df(x)|_{X_1}$ is an isomorphism for all $x \in \widetilde{U}$. For each $x \in \widetilde{U}$, we have

$$DG(x)(X_1) = Df(x)(X_1) \times P(X_1) = Df(x)(X_1) \times \{0\}.$$

It follows that $DG(x)(X_1) = Y \times \{0\}$ and the restriction $DG(x)|_{X_1} : X_1 \to Y \times \{0\}$ is an isomophism for all $x \in \widetilde{U}$.

(2) $\implies$ (3). Define $F = G^{-1}$. Then F is a C^k diffeomorphism from $W \times V$ to $F(W \times V) = \widetilde{U}$. Also, for all $(y, v) \in W \times V$, we have

$$G(F(y, v)) = (f(F(y, v)), P(F(y, v))) = (y, v).$$

This implies that $f(F(y,v)) = y$ for all $(y,v) \in W \times V$. Moreover, for all $(w,v) \in W \times V$, we know $DF(w,v) = [DG(F(y,v))]^{-1}$. Therefore, $DF(w,v)(Y \times \{0\}) = X_1$ and the restriction $DF(w,v)|_{Y\times\{0\}} : Y \times \{0\} \to X$ is an isomophism.

(3) $\implies$ (4). Define $g : W \to X$ by $g(y) = F(y, Px_0)$. It then follows that $g \in C^k(W;X)$, $g(W) \subset U$, and $f(g(y)) = y$ for all $y \in W$.

(4) $\implies$ (5). Since $f(g(y)) = y$ for all $y \in W$ and $y_0 \in W$ we have

$$D(f \circ g)(y_0) = Df(g(y_0))Dg(y_0) = Df(x_0)Dg(y_0) = I_Y$$

This implies that $Df(x_0) \in \mathcal{L}_R(X;Y)$. Recall that the set of right invertible linear maps are open, so we can select open set $x_0 \in \widetilde{U} \subset U$ such that $Df(x) \in \mathcal{L}_R(X;Y)$ for all $x \in \widetilde{U}$. This is equivalent to $Df(x) \in \mathcal{L}(X;Y)$ being surjective with $\ker Df(x)$ closed and complemented in X for all $x \in \widetilde{U}$.

(5) $\implies$ (1). Trivially true since $x_0 \in \widetilde{U}$. □

Theorem (Constant rank theorem)**.** Let X and Y be nontrivial Banach spaces, and suppose that $X_0, X_1 \subset X$ and $Y_0, Y_1 \subset Y$ are nontrivial closed subspaces such that $X = X_0 \oplus X_1$ and $Y = Y_0 \oplus Y_1$. Write P_{X_i} and P_{Y_i} for the projectors onto X_i and Y_i, respectively, for $i = 0, 1$. Let $U \subset X$ be open and $x_0 \in U$. Suppose that $f \in C^k(U;Y)$ for $1 \le k \le \infty$ is such that $\ker Df(x_0) = X_0$ and $\operatorname{range} Df(x_0) = Y_0$, and write $y_0 = f(x_0)$. Further suppose that f satisfies the "constant rank condition":

$$Df(x)(X_1) = \operatorname{range} Df(x) \text{ for all } x \in U.$$

Then there exist open sets $x_0 \in \widetilde{U} \subset U$ and $y_0 \in S \subset Y$ such that $f(\widetilde{U}) \subset S$ and the following hold.

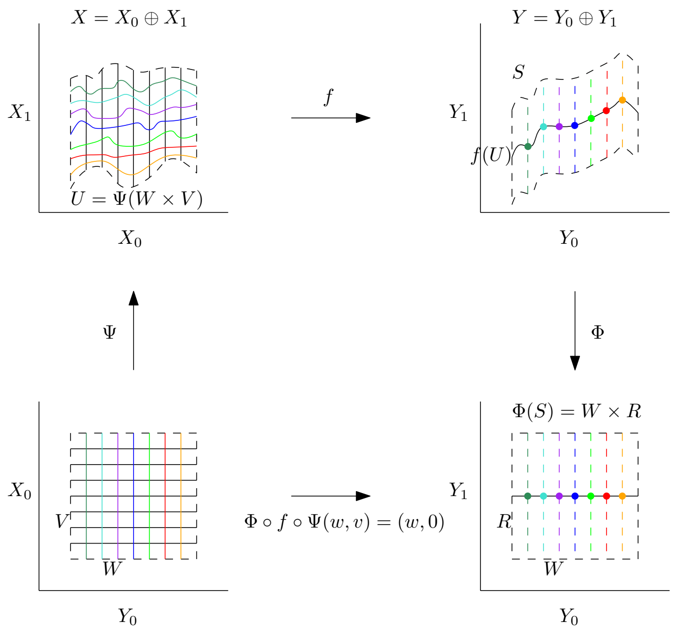
1. There exist open sets $P_{Y_0}y_0 \in W \subset Y_0$ and $P_{X_0}x_0 \in V \subset X_0$ and a C^k diffeomorphism $\Psi : W \times V \to \widetilde{U}$ with $\Psi(P_{Y_0}y_0, P_{X_0}x_0) = x_0$.
2. There exists a C^k diffeomorphism $\Phi : S \to \Phi(S) \subset Y_0 \times Y_1$ with and $\Phi(y_0) = (P_{Y_0}y_0, 0)$.
3. We have that
$$\Phi \circ f \circ \Psi(w,v) = (w,0) \text{ for all } (w,v) \in W \times V.$$

Prove the constant rank theorem by arguing as follows.

NOTE: the final forms of the objects you're after won't appear clearly until step 4. For instance, don't assume that the W from step 1 is the W you're after; you may need to shrink it.

1. Set $w_0 = P_{Y_0}y_0 \in Y_0$ and $v_0 = P_{X_0}x_0 \in X_0$, and define the maps $f_i = P_{Y_i} \circ f$ for $i = 0, 1$. Prove that there exist open sets $x_0 \in \widetilde{U} \subset U$, $w_0 \in W \subset Y_0$, and $v_0 \in V \subset X_0$ with V connected, and a C^k diffeomorphism $\Psi : W \times V \to \Psi(W \times V) = \widetilde{U}$ such that $\Psi(w_0, v_0) = x_0$,
$$f_0(\Psi(w,v)) = w \text{ for all } (w,v) \in W \times V,$$

Depiction of the Constant Rank Theorem



and

$$D\Psi(w, v)|_{Y_0 \times \{0\}} \to X_1 \text{ is an isomorphism for all } (w, v) \in W \times V.$$

2. Define $F : W \times V \to Y$ via $F = f \circ \Psi$. Prove the following for all $(w, v) \in W \times V$.
   (a) $F(w, v) = w + \eta(w, v)$ for some $\eta \in C^k(W \times V; Y_1)$.
   (b) $\operatorname{range} Df(\Psi(w, v)) = \operatorname{range} DF(w, v) = DF(w, v)(Y_0 \times \{0\})$.
   (c) The map $T \in \mathcal{L}(Y; Y_0 \times \{0\})$ defined by $Th = (P_{Y_0}h, 0)$ is a left-inverse of the restriction $DF(w, v)|_{Y_0 \times \{0\}}$.
   (d) $DF(w, v)|_{Y_0 \times \{0\}}$ is injective with closed and complemented range given by

$$\operatorname{range} DF(w, v) = \operatorname{range} Df(\Psi(w, v)).$$

3. Prove that $DF(w, v) \circ T = I$ on $DF(w, v)(Y_0 \times \{0\}) = \operatorname{range} DF(w, v) \subset Y$ for $(w, v) \in W \times V$. Use this to deduce that $D_2\eta(w, v) = 0$ and hence that $D_2F(w, v) = 0$ for $(w, v) \in W \times V$. In turn, deduce that

$$F(w, v) = F(w, v_0) \text{ for all } (w, v) \in W \times V.$$

4. Define $G : W \to Y$ via $G(w) = F(w, v_0)$. Apply the left inverse function theorem to complete the proof.

Proof. 1. Define a map $g : U \to Y_0 \times X_0$ by

$$g(x) = (P_{Y_0}(f(x)), P_{X_0}x)\,.$$

Then, the derivative of g at $x \in U$ is

$$Dg(x)h = (P_{Y_0} \circ Df(x)(h), P_{X_0}h).$$

Claim that $Dg(x_0)$ is an isomorphism. For surjectivity, consider $(y, x) \in Y_0 \times X_0$. Since $\operatorname{range} Df(x_0) = Y_0$, there is $v \in X$ such that $Df(x_0)(v) = y$. Since $\ker Df(x_0) = X_0$ and $X_0 \oplus X_1 = X$, we can assume WLOG that $v \in X_1$. Now, let $h = v + x$. Then,

$$Dg(x_0)h = (P_{Y_0} \circ Df(x_0)(v), P_{X_0}x) = (y, x).$$

Therefore, $Dg(x_0)$ is surjective. For injective, consider $Dg(x_0)h = (0, 0)$ for some $h \in X$. It follows that $P_{X_0}h = 0$ and $Df(x_0)h = Df(x_0)(P_{X_1}h) = 0$. This implies $P_{X_1}h \in X_1 \cap \ker Df(x_0) = X_1 \cap X_0$. Therefore, $h = 0$ and $Dg(x_0)$ is indeed an isomorphism.

Now, the inverse function theorem gives bounded and open sets $x_0 \in U' \subset U$ such that $g : U' \to g(U')$ is a bi-Lipschitz homeomorphism, $Dg(x)$ an isomorphism for all $x \in U'$, and g^{-1} differentiable on $g(U')$ with

$$Dg^{-1}(w, v) = [Dg(g^{-1}(w, v))]^{-1}.$$

Further let $w_0 \in W \subset Y_0$ and $v_0 \in V \subset X_0$ be small open balls such that $W \times V \subset g(U')$, $\Psi : W \times V \to X$ be the inverse of g, and $\widetilde{U} = g(W \times V)$. Now $x_0 \in \widetilde{U}$ since $g(x_0) = (w_0, v_0)$, $v_0 \in V \subset X_0$ connected, and $\Psi : W \times V \to \widetilde{U}$ is a C^k diffeomorphism. Additionally, for all $(w, v) \in W \times V \subset Y_0 \times X_0$, we have let $\Psi(w, v) = x$ for some $x \in \widetilde{U}$. Then, we know $P_{Y_0} \circ f(x) = w$. Therefore,

$$f_0(\Psi(w, v)) = w$$

for all $(w, v) \in W \times V$.

To show $D\Psi(w, v)|_{Y_0 \times \{0\}} \to X_1$ is an isomorphism for all $(w, v) \in W \times V$, note that $D\Psi(w, v)$ is already an isomorphism and thus injective. Now for any $u \in X_1$, we have

$$Dg(\Psi(w, v))u = (P_{Y_0} \circ Df(\Psi(w, v))(u), P_{X_0}u).$$

Note that $P_{X_0}u = 0$ since $u \in X_1$ and $P_{Y_0} \circ Df(\Psi(w, v))(u) \in Y_0$. Also, $Dg(\Psi(w, v))$ is an isomorphism, so

$$D\Psi(w, v)(P_{Y_0} \circ Df(\Psi(w, v))(u), 0) = u.$$

This implies that $D\Psi(w, v)(Y_0 \times \{0\}) = X_1$. Therefore, $D\Psi(w, v)|_{Y_0 \times \{0\}} \to X_1$ is indeed an isomorphism for all $(w, v) \in W \times V$.

2. (a) Define $\eta : W \times V \to Y_1$ by

$$\eta(w, v) = P_{Y_1} \circ f(\Psi(w, v)) = f_1 \circ \Psi(w, v).$$

Then, $\eta \in C^k(W \times V; Y_1)$ and

$$F(w,v) = f(\Psi(w,v)) = f_0(\Psi(w,v)) + f_1(\Psi(w,v)) = w + \eta(w,v),$$

as desired.

(b) First of all, it follows from the chain rule that for all $(w,v) \in W \times V$,

$$DF(w,v) = Df(\Psi(w,v))D\Psi(w,v).$$

Recall from part 1 that $D\Psi(w,v)$ is isomorphism, so range $DF(w,v) =$ range $Df(\Psi(w,v))$. Additionally, we have $Df(\Psi(w,v))(Y_0 \times \{0\}) = X_1$ and note that $Df(x)(X_1) =$ range $Df(x)$ for all $x \in \widetilde{U}$. Therefore,

$$Df(\Psi(w,v))D\Psi(w,v)(Y_0 \times \{0\}) = Df(\Psi(w,v))(X_1) = \operatorname{range} Df(\Psi(w,v)).$$

This implies that

$$\operatorname{range} Df(\Psi(w,v)) = \operatorname{range} DF(w,v) = DF(w,v)(Y_0 \times \{0\}),$$

as desired.

(c) Let $y \in Y_0$. We want to show that $T \circ DF(w,v)(y,0) = (y,0)$. Again, note that

$$DF(w,v) = Df(\Psi(w,v))D\Psi(w,v).$$

Suppose $D\Psi(w,v)(y,0) = h$. As shown in part 1, we have

$$Dg(x)h = (P_{Y_0} \circ Df(x)(h), P_{X_0}h).$$

for all $x \in U$. It follows that $P_{Y_0} \circ Df(\Psi(w,v))h = y$ and $P_{X_0}h = 0$. Therefore,

$$T \circ DF(w,v)(y,0) = T \circ Df(\Psi(w,v))h = (P_{Y_0} \circ Df(\Psi(w,v))h, 0) = (y,0).$$

This implies that $T \in \mathcal{L}(Y; Y_0 \times \{0\})$ defined by $Th = (P_{Y_0}h, 0)$ is a left inverse of $DF(w,v)$ restricted on $Y_0 \times \{0\}$.

(d) Since $DF(w,v)|_{Y_0 \times \{0\}}$ has a left inverse, $DF(w,v)|_{Y_0 \times \{0\}}$ is injective with its range closed and complemented. In addition, from part (b), we know that range $DF(w,v) =$ range $Df(\Psi(w,v))$, as desired.

3. Since $DF(w,v)|_{Y_0 \times \{0\}}$ is injective with closed and complemented range given by

$$\operatorname{range} DF(w,v) = \operatorname{range} Df(\Psi(w,v)),$$

$DF(w,v) : Y_0 \times \{0\} \to$ range $DF(w,v)$ is an isomorphism. In particular, it is invertible. Since T is the a left-inverse, T must also be the right inverse. Therefore, $DF(w,v) \circ T = I$ on $DF(w,v)(Y_0 \times \{0\}) =$ range $DF(w,v) \subset Y$ for all $(w,v) \in W \times V$.

Next we show $D_2\eta(w,v)=0$. Let $h \in X_0$ be arbitrary. By the definition of η, we have

$$\begin{aligned} D_2\eta(h) &= P_{Y_1} \circ Df(\Psi(w,v))D_2\Psi(w,v)h \\ &= P_{Y_1} \circ Df(\Psi(w,v))D\Psi(0,h) \\ &= P_{Y_1} \circ DF(w,v)(0,h). \end{aligned}$$

Now, using the fact that $DF(w,v) \circ T = I$, we have

$$\begin{aligned} DF(w,v)(0,h) &= DF(w,v) \circ T \circ DF(w,v)(0,h) \\ &= DF(w,v)(P_{Y_0} \circ DF(w,v)(0,h), 0). \end{aligned}$$

Claim that $P_{Y_0} \circ DF(w,v)(0,h) = 0$. Indeed, let $D\Psi(w,v)(0,h) = u$. Note that $D\Psi(w,v) = [Dg(\Psi(w,v))]^{-1}$ and $Dg(x)k = (P_{Y_0} \circ Df(x)k, P_{X_0}k)$ for all $x \in U$ and $k \in X$. It follows that $P_{Y_0} \circ Df(\Psi(w,v))u = 0$. Therefore,

$$P_{Y_0} \circ DF(w,v)(0,h) = P_{Y_0} \circ Df(\Psi(w,v))D\Psi(w,v)(0,h) = P_{Y_0} \circ Df(\Psi(w,v))u = 0.$$

Hence,

$$D_2\eta(h) = P_{Y_1} \circ DF(w,v)(0,h) = P_{Y_1} \circ DF(w,v)(0,0) = 0.$$

This implies that $D_2\eta(w,v) = 0$, as desired.

Now, by the equality in part 2(a), we have

$$D_2F(w,v) = D_2\eta(w,v) = 0.$$

for all $(w,v) \in W \times V$. Now by the mean value theorem,

$$\|F(w,v) - F(w,v_0)\| \le \|v - v_0\| \sup_{h \in V} \|D_2F(w,h)\| = 0.$$

Therefore, $F(w,v) = F(w,v_0)$ for all $(w,v) \in W \times V$.

4. Define $G : W \to Y$ by $G(w) = F(w,v_0)$. Claim that $DG(w_0) \in \mathcal{L}(Y_0; Y)$ is injective with range $DG(w_0)$ closed and complemented. Indeed, for any $Y_0 \in X$, we have

$$DG(w)h = D_1F(w,v_0)h = DF(w,v_0)(h,0).$$

Note that in part 2(d), we have shown that $DF(w,v)|_{Y_0 \times \{0\}}$ is injective with closed and complemented range. Therefore, $DG(w_0)$ is also injective with closed and complemented range. In fact,

$$\operatorname{range} DG(w_0) = DF(w_0,v_0)(Y_0 \times \{0\}) = \operatorname{range} Df(\Psi(w_0,v_0)) = \operatorname{range} Df(x_0) = Y_0.$$

Note that $Y_0 \oplus Y_1 = Y$. Now, by the left inverse function theorem, thre exists open sets $w_0 \in \widetilde{W} \subset W$, $0 \in R \subset Y_1$, and $S \subset Y$ and a map $\Phi \in C^k(S \to Y_0 \times Y_1)$ such that Φ is a C^k diffeomorphism from S to $\Phi(S) = \widetilde{W} \times R \subset Y_0 \times Y_1$, $G(\widetilde{W}) \subset S$, and

$$\Phi(G(w)) = (w, 0)$$

for all $w \in \widetilde{W}$. In particular,

$$\Phi(y_0) = \Phi(F(w_0, v_0)) = \Phi(G(w_0)) = (w_0, 0),$$

and

$$\Phi \circ f \circ \Psi(w, v) = \Psi(F(w, v_0)) = \Psi(G(w)) = (w, 0)$$

for all $w \in \widetilde{W}$. This completes the proof with $\widetilde{W}$ being the desired W in the problem statement.

□

Below is a naming sanity check for the constant rank theorem.

Theorem. Let $X = \mathbb{F}^{n+r}$ and $Y = \mathbb{F}^{m+r}$ for $1 \leq m, n, r$. Suppose that $U \subset X$ is open with $x_0 \in U$, and $f \in C^k(U; Y)$ for $1 \leq k \leq \infty$. Write $\ker Df(x_0) = X_0$ and $\operatorname{range} Df(x_0) = Y_0$. In finite dimensions all subspaces are closed and complemented, so we write $X = X_0 \oplus X_1$ and $Y = Y_0 \oplus Y_1$. Prove that the following are equivalent.

1. There exists an open set $x_0 \in \widetilde{U} \subset U$ such that $\operatorname{rank} Df(x) = r$ for all $x \in \widetilde{U}$.
2. We have that $\dim X_1 = \dim Y_0 = r$, and there exists an open set $x_0 \in \widetilde{U} \subset U$ such that

$$Df(x)(X_1) = \operatorname{range} Df(x) \text{ for all } x \in \widetilde{U}.$$

NOTE: this equivalence is where the constant rank theorem gets its name.

Proof. (1) $\implies$ (2). Suppose there exists an open set $x_0 \in U' \subset U$ such that $\operatorname{rank} Df(x) = r$ for all $x \in U'$. Note that $\ker Df(x_0) = X_0$, $\operatorname{range} Df(x_0) = Y_0$. This implies that $\dim X_1 = \dim Y_0 = \operatorname{rank} Df(x_0) = r$. Additionally, this implies that $Df(x_0)$ is injective with closed range when restricted to X_1. Since the set of injective maps with closed range is open and $x_0 \in U'$ is open, we can find open set $x_0 \in \widetilde{U} \subset U'$ so small that $Df(x)$ is injective with closed range when restricted to X_1 for all $x \in \widetilde{U}$. It follows that for all $x \in \widetilde{U}$, we have

$$\dim Df(x)(X_1) \geq \dim X_1 = r.$$

However, note that $\operatorname{rank} Df(x) = \dim \operatorname{range} Df(x) = r$ and $Df(x)(X_1) \subset \operatorname{range} Df(x)$ for all $x \in \widetilde{U}$. Therefore,

$$Df(x)(X_1) = \operatorname{range} Df(x)$$

for all $x \in \widetilde{U}$.

(2) $\implies$ (1). Suppose $\dim X_1 = \dim Y_0 = r$ and there exists an open set $x_0 \in U' \subset U$ such that

$$Df(x)(X_1) = \operatorname{range} Df(x)$$

for all $x \in U'$. Let $y_0 = f(x_0)$. Then, by constant rank theorem, there exists open sets $x_0 \in \widetilde{U} \subset U'$, $P_{Y_0} y_0 \in W \subset Y_0$, and $P_{X_0} x_0 \in V \subset X_0$, and a C^k diffeomorphism $\Psi : W \times V \to \widetilde{U}$.

Additionally, let $F : W \times V \to Y$ by $F = f \circ \Psi$. Then for all $x \in \widetilde{U}$, we know $x = \Psi(w, v)$ for some $(w, v) \in W \times V$. Also, $\operatorname{range} Df(x) = DF(w, v)(Y_0 \times \{0\})$ with $DF(w, v)|_{Y_0 \times \{0\}}$ an isomorphism from $Y_0 \times \{0\}$ to $DF(w, v)(Y_0 \times \{0\})$. It follows that

$$\operatorname{rank} Df(x) = \dim \operatorname{range} Df(x) = \dim DF(w, v)|_{Y \times \{0\}} = \dim Y_0 = r.$$

for all $x \in \widetilde{U}$. This completes the proof.

□

We define now a nonlinear version of linear dependence.

Definition. Assume that $m, n \in \mathbb{N}^+$. Let $\emptyset \neq U \subset \mathbb{R}^n$, and suppose that $f_i : U \to \mathbb{R}$ for $i = 1, \dots, m$. Write $f : U \to \mathbb{R}^m$ via $f = (f_1, \dots, f_m)$. We say that $\{f_1, \dots, f_m\}$ are functionally dependent in U if there exists an open set $V \subset \mathbb{R}^m$ with $f(U) \subset V$ and a function $g \in C^1(V; \mathbb{R})$ such that

$$g(f_1(x), \dots, f_m(x)) = 0 \text{ for all } x \in U$$

and $\nabla g(y) \neq 0$ for all $y \in V$.

Theorem. Let $\emptyset \neq U \subset \mathbb{R}^n$ and $f \in C^1(U; \mathbb{R}^m)$.

1. Prove that if $\{f_1, \dots, f_m\}$ are functionally dependent on U, then for each $x_0 \in U$ there exists an open set $V \subset U$ with $x_0 \in V$, a function $h \in C^1(W; \mathbb{R})$, where $W \subset \mathbb{R}^{m-1}$ is an open set, and an index $i \in \{1, \dots, m\}$ such that

   $$f_i(x) = h(f_1(x), \dots, f_{i-1}(x), f_{i+1}(x), \dots, f_m(x)) \text{ for all } x \in V.$$

   In other words, at least one of the functions can be written in terms of the other functions.

2. Prove that if $m \leq n$ and $\{f_1, \dots, f_m\}$ are functionally dependent on U, then

   $$\operatorname{rank}(Df(x)) \leq m - 1 \text{ for all } x \in U.$$

3. Suppose that $1 \leq k < \min\{m, n\}$ and that Df is of constant rank k on U. Prove that for each $x_0 \in U$ there exists an open set $Z \subset U$ such that $x_0 \in Z$ and $\{f_1, \dots, f_m\}$ are functionally dependent in Z.

Proof. 1. Let $x_0 \in U$ and $y_0 = f(x_0)$. Then, we have

$$\nabla g(y_0) = (D_1 g(y_0), \dots, D_m g(y_0)) \neq 0.$$

It follows that there exists an index $i \in \{1, \dots, m\}$ such that $D_i g(y_0) \neq 0$. WLOG assume $i = 1$. Now, by the implicit function theorem, there exists open sets $(f_2(x_0), \dots, f_m(x_0)) \in W \subset \mathbb{R}^{m-1}$, $f_1(x_0) \in \widetilde{V} \subset \mathbb{R}$, and $0 \in S \subset \mathbb{R}$ and $h \in C^1(W; \mathbb{R})$ such that

$$h(f_2(x_0), \dots, f_m(x_0)) = f_1(x_0)$$

and $(y_2, \ldots, y_m, h(y_2, \ldots, y_m, z)) \in W \times \widetilde{V}$ for all $(y_2, \ldots, y_m, z) \in W \times S$. Additionally,

$$g(y_2, \ldots, y_m, h(y_2, \ldots, y_m, z)) = z$$

for all $(y_2, \ldots, y_m, z) \in W \times S$. Also, if $(y_1, y_2, \ldots, y_m) \in \widetilde{V} \times W$ such that

$$g(y_1, y_2, \ldots, y_m) = z$$

for some $z \in S$, then $y_1 = h(y_2, \ldots, y_m)$.

Since $f_i \in C^1(U; \mathbb{R})$ for all $i = 1, \ldots, m$, we can select open set $x_0 \in V \subset U$ so small that $f_1(x) \in \widetilde{V}$ and $(f_2(x), \ldots, f_m(x)) \in W$ for all $x \in V$. It follows that

$$g(f_1(x), \ldots, f_m(x)) = 0$$

for all $x \in V$. This implies that $f_1(x) = h(f_2(x), \ldots, f_m(x))$ as argued above.

2. Since $g(f_1(x), \ldots, f_m(x)) = g(f(x)) = 0$ for all $x \in U$, we have

$$\nabla g(f(x)) \circ Df(x) = 0.$$

Notice that $f(U) \subset V$ and $\nabla g(y) \neq 0$ for all $y \in V$. It then follows that $\nabla g(f(x))(\mathbb{R}^m) \neq 0$ for all $x \in U$. Suppose for contradiction that $\operatorname{rank} Df(x) = m$ for some $x_0 \in U$. Then, $Df(x_0)(\mathbb{R}^n) = \mathbb{R}^m$. However, this will imply that $\nabla g(f(x_0)) \circ Df(x_0) \neq 0$, a contradiction. Therefore,

$$\operatorname{rank} Df(x) \leq m - 1$$

for all $x \in \widetilde{U}$.

3. Suppose that $1 \leq k < \min\{m, n\}$ and that $\operatorname{rank} Df(x) = k$ for all $x \in U$. Fix $x_0 \in U$ and write $\ker Df(x_0) = X_0$ and $\operatorname{range} Df(x_0) = Y_0$, where X_0 and Y_0 are complemented with $X_0 \oplus X_1 = \mathbb{R}^n$ and $Y_0 \oplus Y_1 = \mathbb{R}^m$. By the previous sanity check, we know that $\dim X_1 = \dim Y_0 = k$ and there exists and open set $x_0 \in \widetilde{U} \subset U$ such that

$$Df(x)(X_1) = \operatorname{range} Df(x)$$

for all $x \in \widetilde{U}$.

Note that $k < \min\{m, n\}$, so X_i and Y_i are nontrivial for $i = 0, 1$. Set $y_0 = f(x_0)$. By constant rank theorem, there exists open sets $x_0 \in Z \subset U$ and $y_0 \in S \subset Y$, $P_{Y_0} y_0 \in W \subset Y_0$, and $P_{X_0} x_0 \in V \subset X_0$, and C^k diffeomorphisms $\Psi : W \times V \to Z$ with $\Psi(P_{Y_0} y_0, P_{X_0} x_0) = x_0$, $\Phi : S \to Y_0 \times Y_1$ with $\Phi(y_0) = (P_{Y_0} y_0, 0)$, and

$$\Phi \circ f \circ \Psi(w, v) = (w, 0)$$

for all $w \in W$. Now define $g : S \to \mathbb{R}$ by

$$g(y) = \pi_{k+1} \circ \Phi(y),$$

where π_{k+1} is projection onto the $(k+1)$-th index. Notice first for all $y \in S$, we have

$$\nabla g(y) = \pi_{k+1} \circ D\Phi(y).$$

Since $D\Phi(y)$ is an isomorphism and π_{k+1} has nonzero range, $\nabla g(y) \neq 0$ for all $y \in S$. Additionally, for each $x \in Z$, there exists $(w, v) = \Psi^{-1}(x) \in W \times V$ Therefore, for all $x \in Z$,

$$g(f(x)) = \pi_{k+1} \circ \Phi \circ f \circ \Psi(\Psi^{-1}(x)) = \pi_{k+1}(w, 0) = 0.$$

This shows that $\{f_1, \ldots, f_m\}$ are functionally dependent in Z.

□

3 Measure and integration

3.1 Introduction to abstrct measure theory

3.1.1 Basic definitions

Definition. Let X be a set.

1. An **algebra** on X is $\mathfrak{A} \subset \mathcal{P}(X)$ such that
   (a) $\emptyset \in \mathfrak{A}$.
   (b) $E \in \mathfrak{A}$ implies $E^c \in \mathfrak{A}$.
   (c) $E, F \in \mathfrak{A}$ implies $E \cup F \in \mathfrak{A}$.
2. A **σ-algebra** is an algebra $\mathfrak{M} \subset \mathcal{P}(X)$ such that if $E_k \in \mathfrak{M}$ for all $k \in \mathbb{N}$, then $\bigcup_{k=0}^{\infty} E_k \in \mathfrak{M}$.
3. A pair $(X, \mathfrak{M})$ with $\mathfrak{M}$ a σ-algebra on X is called a **measurable space**.

Theorem. Let X be a set.

1. Suppose $A \neq \emptyset$ is a set and $\mathfrak{M}_\alpha$ is σ-algebra for each $\alpha \in A$, then $\mathfrak{M} = \bigcap_{\alpha \in A} \mathfrak{M}_\alpha$ is a σ-algbera on X.
2. Suppose $F \subset \mathcal{P}(X)$, there is unique smallest σ-algebra $\mathfrak{M}$ on X such that $F \subset \mathfrak{M}$. Write $\mathfrak{M} = \sigma(F)$ and call this the σ-algebra generated by F.

Theorem. Let X and Y be sets and $f : X \to Y$.

1. Suppose $\mathfrak{M}$ is a σ-algebra on X and set
$$\mathfrak{N} = \left\{ E \subset Y : f^{-1}(E) \in \mathfrak{M} \right\}.$$
Then, $\mathfrak{N}$ is a σ-algebra on Y. Call this the **push-forward** of $\mathfrak{M}$ by f.
2. Suppose $\mathfrak{N}$ is a σ-algebra on Y and set
$$\mathfrak{M} = \left\{ f^{-1}(E) : E \in \mathfrak{N} \right\}.$$
Then, $\mathfrak{M}$ is a σ-algebra on X. Call this the **pull-back** of $\mathfrak{N}$ by f.

Definition. Let $A \neq \emptyset$ be a set.

1. Let Y be a set and X_α be sets with σ-algebra $\mathfrak{M}_\alpha$ for all $\alpha \in A$. Suppose $g_\alpha : X_\alpha \to Y$ for all $\alpha \in A$. Define
$$\sigma\left(\left\{ E \subset Y : g_\alpha^{-1}(E) \in \mathfrak{M}_\alpha \text{ for all } \alpha \in A \right\} \right)$$
to be the **push-forward** of $\{g_\alpha\}_{\alpha \in A}$.
2. Let X be a set and Y_α be sets with σ-algebra $\mathfrak{N}_\alpha$ for all $\alpha \in A$. Suppose $f_\alpha : X \to Y_\alpha$ for all $\alpha \in A$. Define
$$\sigma\left(\left\{ f_\alpha^{-1}(E) : E \in \mathfrak{N}_\alpha \text{ for some } \alpha \in A \right\} \right)$$
to be the **pull-back** of $\{f_\alpha\}_{\alpha \in A}$.

Definition. Let $A \neq \emptyset$ be a set and X_α be sets with σ-algebra $\mathfrak{M}_\alpha$ for all $\alpha \in A$. Then on the set $X = \prod_\alpha X_\alpha$ we define the **product σ-algbera** $\bigotimes_\alpha \mathfrak{M}_\alpha$ to be the pull-back of projection maps $\pi_\alpha : X \to X_\alpha$.

Theorem. Let $A \neq \emptyset$ be a set and X_α with σ-algebra $\mathfrak{M}_\alpha$ for all $\alpha \in A$. Let $X = \prod_\alpha X_\alpha$ and define

$$\mathcal{R} = \left\{ \prod_\alpha M_\alpha : M_\alpha \in \mathfrak{M}_\alpha \right\}.$$

Then,

1. $\bigotimes_\alpha \mathfrak{M}_\alpha \subset \sigma(\mathcal{R})$. If A countable then $\sigma(\mathcal{R}) = \bigotimes_\alpha \mathfrak{M}_\alpha$.
2. Suppose $\mathfrak{M}_\alpha = \sigma(\mathcal{E}_\alpha)$ for all $\alpha \in A$ and let

$$\mathcal{E} = \left\{ \pi_\alpha^{-1}(E) : E \in \mathcal{E}_\alpha \text{ for some } \mathcal{E}_\alpha \right\}.$$

Then $\bigotimes_\alpha \mathfrak{M}_\alpha = \sigma(\mathcal{E})$. Moreover, if A is countable and $X_\alpha \in \mathcal{E}_\alpha$ for all $\alpha \in A$, then $\bigotimes_\alpha \mathfrak{M}_\alpha$ is generated by $\mathcal{F} = \{\prod_\alpha E_\alpha : E_\alpha \in \mathcal{E}_\alpha\}$

Proof. 1. For $E \in \mathfrak{M}_\alpha$, we have $\pi_\alpha^{-1}(E) = \prod_\beta S_\beta$, where

$$S_\beta = \begin{cases} E & (\beta = \alpha), \\ X_\beta & (\beta \neq \alpha). \end{cases}$$

Then,

$$\left\{ \pi_\alpha^{-1}(M_\alpha) : M_\alpha \in \mathfrak{M}_\alpha \right\} \subset \left\{ \prod_\beta M_\beta : M_\beta \in \mathfrak{M}_\beta \right\} = \mathcal{R}.$$

This implies that $\bigotimes_\alpha \mathfrak{M}_\alpha \subset \sigma(\mathcal{R})$.

On the other hand, if A is countable, then

$$\prod_\alpha M_\alpha = \bigcap_\alpha \pi_\alpha^{-1}(M_\alpha) \in \bigotimes_\alpha \mathfrak{M}_\alpha$$

whenever $M_\alpha \in \mathfrak{M}_\alpha$ for all $\alpha \in A$. This implies that $\sigma(\mathcal{R}) \subset \bigotimes_\alpha \mathfrak{M}_\alpha$.

2. It is clear that $\sigma(\mathcal{E}) \subset \bigotimes_\alpha \mathfrak{M}_\alpha$. On the other hand, for each $\alpha \in A$, let

$$\mathfrak{N}_\alpha = \left\{ E \subset X_\alpha : \pi_\alpha^{-1}(E) \in \sigma(\mathcal{E}) \right\}$$

be the push-forward of $\sigma(\mathcal{E})$ to X_α by π_α. It is clear that $\mathcal{E}_\alpha \subset \mathfrak{N}_\alpha$. This implies $\mathfrak{M}_\alpha = \sigma(\mathcal{E}) \subset \mathfrak{N}_\alpha$. In particular, $\pi_\alpha^{-1}(E) \in \sigma(\mathcal{E})$ for all $E \in \mathfrak{M}_\alpha$. This implies that $\bigotimes_\alpha \mathfrak{M}_\alpha \subset \sigma(\mathcal{E})$.

Now, assume A countable and $X_\alpha \in \mathcal{E}_\alpha$ for all $\alpha \in A$. Then let $E \in \mathfrak{M}_\alpha$ for some $\alpha \in A$. We have $\pi_\alpha^{-1}(E) = \prod_\beta S_\beta$, where

$$S_\beta = \begin{cases} E & (\beta = \alpha), \\ X_\beta & (\beta \neq \alpha). \end{cases}$$

Therefore, $\sigma(\mathcal{E}) \subset \sigma(\mathcal{F})$.

On the other hand, since A is countable, we have

$$\prod_\alpha E_\alpha = \bigcap_\alpha \pi_\alpha^{-1}(E_\alpha) \in \sigma(\mathcal{E}).$$

This implies that $\sigma(\mathcal{F}) \subset \sigma(\mathcal{E})$ and the proof is complete.

$\square$

Corollary. If $\mathfrak{M}_i$ is σ-algebra for $i = 1, 2, 3$, then

$$\mathfrak{M}_1 \otimes (\mathfrak{M}_2 \otimes \mathfrak{M}_3) = (\mathfrak{M}_1 \otimes \mathfrak{M}_2) \otimes \mathfrak{M}_3 = \mathfrak{M}_1 \otimes \mathfrak{M}_2 \otimes \mathfrak{M}_3,$$

since they are all generated by

$$\{M_1 \times (M_2 \times M_3)\} = \{(M_1 \times M_2) \times M_3\} = \{M_1 \times M_2 \times M_3\}.$$

Theorem. Let $X_1, \dots, X_n$ be metric spaces and $X = \prod_{i=1}^n X_i$ be equipped with the ususal metric. Then, $\bigotimes_{i=1}^n \mathfrak{B}_{X_i} \subset \mathfrak{B}_X$. However, if each X_i is separable, then $\mathfrak{B}_X = \bigotimes_{i=1}^n \mathfrak{B}_{X_i}$.

Proof. We know by the previous theorem that $\bigotimes_{i=1}^n \mathfrak{B}_{X_i}$ is generated by $\{\prod_i U_i : U_i \subset X_i \text{ open}\}$. However, $\prod_i U_i$ is open in X. Therefore, $\bigotimes_{i=1}^n \mathfrak{B}_{X_i} \subset \mathfrak{B}_X$.

Suppose now each X_i is separable and let $D_i \subset X_i$ be countable and dense. Consider

$$\mathcal{E}_i = \left\{ B(x_i, r) : \ X_i \in D_i,\ r = \infty \text{ or } r \in \mathbb{Q}^+ \right\},$$

which is countable and $\sigma(\mathcal{E}_i) = \mathfrak{B}_{X_i}$ since every open set in X_i is countable union of elements in $\mathcal{E}_i$. Similarly, $\mathfrak{B}_X$ is generated by $\{\prod_i E_i : E_i \in \mathcal{E}_i\}$. But item 2 from the previous theorem implies that $\bigotimes_{i=1}^n \mathfrak{B}_{X_i}$ is generated by the same set. Therefore, $\bigotimes_{i=1}^n \mathfrak{B}_{X_i} = \mathfrak{B}_X$.

$\square$

Remark. The above theorem is not true in general if X_i is not separable for some i.

Definition. Let X be a metric space. Define

$$F_\sigma(X) = \left\{ \bigcup_{k=0}^\infty C_k : \ C_k \subset X \text{ closed} \right\},$$

$$G_\delta(X) = \left\{ \bigcap_{k=0}^\infty U_k : \ U_k \subset X \text{ open} \right\}.$$

Note that $F_\sigma(X) \subset \mathfrak{B}_X$ and $G_\delta(X) \subset \mathfrak{B}_X$.

Theorem. Let X be a metric space. Then the following holds:

1. F_σ and G_δ are both closed under finite union and intersection.
2. If $C \subset X$ is closed, then $C \in G_\delta$. If $U \subset X$ is open, then $U \in F_\sigma$.
3. Suppose X is σ-compact, that is, $X = \bigcup_{n=0}^\infty K_n$ for $K_n \subset X$ compact, then each $F \in F_\sigma$ is also σ-compact. In particular, all open sets are σ-compact.

Theorem. Let X and Y be metric spaces and $f : X \to Y$ be continuous. Then the following holds:

1. $E \in F_\sigma(Y)$ implies that $f^{-1}(E) \in F_\sigma(X)$, and $E \in G_\delta(Y)$ implies that $f^{-1}(E) \in G_\delta(X)$.
2. If $E \in \mathfrak{B}(Y)$, then $f^{-1}(E) \in \mathfrak{B}(X)$.

Theorem. Let X and Y be metric spaces with X σ-compact. Then,

1. If $E \in F_\sigma(X)$ and $f : E \to Y$ is continuous, then $f(E) \in F_\sigma(Y)$ and σ-compact.
2. If $f : X \to Y$ is a continuous injection, then $E \in \mathfrak{B}(X)$ implies $f(E) \in \mathfrak{B}(Y)$.

Corollary. Let $\emptyset \neq X \subset Y$ for Y a metric space. Then $\mathfrak{B}(X) = \mathfrak{B}(Y)_X := \{X \cap E : E \in \mathfrak{B}(Y)\}$.

Proof. We know $V \subset X$ open if and only if $V = X \cap U$ for some U open in Y. Therefore,

$$\{V \subset X : V \text{ open in } X\} \subset \mathfrak{B}(Y)_X.$$

This implies that $\mathfrak{B}(X) \subset \mathfrak{B}(Y)_X$.

On the other hand, the inclusion map $I : X \to Y$ is a continuous injection, so if $E \in \mathfrak{B}(Y)$, then $I^{-1}(E) \in \mathfrak{B}(X)$. However, $I^{-1}(E) = E \cap X$. Therefore, $\mathfrak{B}(Y)_X \subset \mathfrak{B}(X)$.

□

3.1.2 Measures

Definition (Measure)**.** Let X be a set with $\mathfrak{M}$ a σ-algebra on X. A **measure** is a map $\mu : \mathfrak{N} \to [0, \infty]$ such that

1. $\mu(\emptyset) = 0$.
2. If $\{E_k\}_{k=0}^\infty \subset \mathfrak{M}$ pairwise disjoint, then $\mu\left(\bigcup_{k=0}^\infty E_k\right) = \sum_{k=0}^\infty \mu(E_k)$.

Such a triple $(X, \mathfrak{M}, \mu)$ is a **measure space**.

Definition. We say $(X, \mathfrak{M}, \mu)$ is **finite** if $\mu(X) < \infty$. We say $(X, \mathfrak{M}, \mu)$ is **σ-finite** if $X = \bigcup_{n=0}^\infty X_n$ for $X_n \in \mathfrak{M}$ and $\mu(X_n) < \infty$.

Theorem. Let $(X, \mathfrak{M}, \mu)$ be a measure space. Then the following holds:

1. If E and F is measurable and $E \subset F$, then $\mu(E) \leq \mu(F)$.

2. If $E_k \in \mathfrak{M}$ for all $k \in \mathbb{N}$, then $\mu\left(\bigcup_{k=0}^{\infty} E_k\right) \leq \sum_{k=0}^{\infty} \mu(E_k)$.

3.1.3 Outer measures and Carathéodory construction

Definition (Outer measure)**.** Let X be a set. An **outer measure** is a map $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that

1. $\mu^*(\emptyset) = 0$.
2. $E \subset F$ implies $\mu^*(E) \leq \mu^*(F)$.
3. If $E_k \subset X$ for all $k \in \mathbb{N}$, then $\mu^*\left(\bigcup_{k=0}^{\infty} E_k\right) \leq \sum_{k=0}^{\infty} \mu^*(E_k)$.

Proposition. Let $\mu_\alpha^* : \mathcal{P}(X) \to [0, \infty]$ be an outer measure for all $\alpha \in A \neq \emptyset$. Then $\lambda : \mathcal{P}(X) \to [0, \infty]$ defined by $\lambda(E) = \sup_{\alpha \in A} \mu_\alpha^*(E)$ is an outer measure.

Proof. 1. $\mu_\alpha^*(\emptyset) = 0$ for all $\alpha \in A$ implies that $\lambda(\emptyset) = 0$.

2. Suppose $E \subset F$, then $\mu_\alpha^*(E) \leq \mu_\alpha^*(F) \leq \lambda(F)$ for all $\alpha \in A$. Take the sup and we obtain $\lambda(E) \leq \lambda(F)$.
3. Let $E_k \subset X$ for each $k \in \mathbb{N}$. Then,

$$\mu_\alpha^*\left(\bigcup_{k=0}^{\infty} E_k\right) \leq \sum_{k=0}^{\infty} \mu_\alpha^*(E_k) \leq \sum_{k=0}^{\infty} \lambda(E_k)$$

   This implies that $\lambda(\bigcup_{k=0}^{\infty} E_k) \leq \sum_{k=0}^{\infty} \lambda(E_k)$.

□

Definition. Let X be a set with outer measure μ^*. Say a set $E \subset X$ is measurable with respect to μ^* if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

for all $A \subset X$.

Theorem (Carathéodory construction)**.** Let X be a set with outer measure μ^*, the following holds.

1. The collection $\mathfrak{M} = \{E \subset X : \ E \text{ measurable}\}$ is a σ-algebra.
2. If $E \subset X$ is such that $\mu^*(E) = 0$, then $E \in \mathfrak{M}$.
3. The restriction $\mu = \mu^*|_{\mathfrak{M}}$ is a measure, and $(X, \mathfrak{M}, \mu)$ is a complete measure space.

Definition (Cover regular)**.** Let μ^* be an outer measure on X. Say μ^* is cover-regular if for any $A \subset X$, there exists $E \in \mathfrak{M}$ such that $A \subset E$ and $\mu^*(A) = \mu(E)$.

Proposition. Let μ^* be an outer measure on X. Then μ^* is outer-regular if and only if for any $A \subset X$, $\mu^*(A) = \inf\{\mu(E) : A \subset E \in \mathfrak{M}\}$. In either case, the inf is a min.

Proposition. Let X be a set with cover-regular outer measure μ^*. Suppose for $n \in \mathbb{N}$, we have $A_n \subset A_{n+1}$. Then,

$$\mu^* \left(\bigcup_{n=0}^{\infty} A_n \right) = \lim_{n \to \infty} \mu^*(A_n).$$

Proof. First note that $\mu^*(A_n) \leq \mu^*(A_{n+1}) \leq \mu^*(A)$, where $A = \bigcup_{n=0}^{\infty} A_n$. Therefore,

$$\lim_{n \to \infty} \mu^*(A_n) \leq \mu^*(A).$$

On the other hand, by cover regularity, there exists $A_n \subset E_n \in \mathfrak{M}$ such that $\mu^*(A_n) = \mu(E_n)$. In particular, $\lim_{n\to\infty} \mu^*(A_n) = \lim_{n\to\infty} \mu(E_n)$. Then,

$$A = \bigcup_{n=0}^{\infty} A_n = \bigcup_{n=0}^{\infty} \bigcap_{k=n}^{\infty} A_k \subset \bigcup_{n=0}^{\infty} \bigcap_{k=n}^{\infty} E_k \in \mathfrak{M},$$

and

$$\mu^*(A) \leq \mu \left(\bigcup_{n=0}^{\infty} \bigcap_{k=n}^{\infty} E_k \right) = \lim_{n \to \infty} \mu \left(\bigcap_{k=n}^{\infty} E_k \right) \leq \lim_{n \to \infty} \mu(E_n) = \lim_{n \to \infty} \mu(A_n),$$

where we have used monotone continuity of **measure**. Therefore,

$$\lim_{n \to \infty} \mu^*(A_n) = \mu^* \left(\bigcup_{n=0}^{\infty} A_n \right).$$

□

3.1.4 Constructing outer measures

Definition. Let X be a set. A gauge on X is a pair $(\mathcal{E}, \gamma)$ where $\mathcal{E} \subset \mathcal{P}(X)$ is such that $\emptyset \in \mathcal{E}$ and $\gamma : \mathcal{E} \to [0, \infty]$ is such that $\gamma(\emptyset) = 0$.

Theorem. Let X be a set and $(\mathcal{E}, \gamma)$ be a gauge on X. Define $\mu^* : \mathcal{P}(X) \to [0, \infty]$ via

$$\mu^*(E) = \inf \left\{ \sum_{n=0}^{\infty} \gamma(E_n) : E \subset \bigcup_{n=0}^{\infty} E_n \text{ and } \{E_n\}_{n=0}^{\infty} \subset \mathcal{E} \right\}.$$

Then μ^* is an outer measure on X and hence generates $(X, \mathfrak{M}, \mu)$, a complete measure space thorugh Carathéodory construction.

Proof. ***** TO-DO ***** □

Theorem. Let (X, d) be a metric space and μ^* be an outer measure on X. Write $\mathfrak{M}$ for the σ-algebra of sets that are measurable with respect to μ^*. Then $\mathfrak{B}(X) \subset \mathfrak{M}$ if and only if μ^* is a metric outer measure.

Proof. ($\implies$). Suppose μ^* is a metric outer measure. Following the hint, we proceed to show that all closed sets are in $\mathfrak{M}$. Let $E \subset X$ be closed and $A \subset X$ be arbitrary. For each $k \in \mathbb{N}$, let

$$F_k = \left\{x \in X : \operatorname{dist}(x, E) > 2^{-k}\right\} \subset E^c.$$

It is easy to see that $F_k \subset F_{k+1}$ for each $k \in \mathbb{N}$. Also, since E is closed, $\operatorname{dist}(x, E) > 0$ for each $x \in E^c$. It follows that $\bigcup_{k=0}^{\infty} F_k = E^c$. Moreover, for any $x \in F_k$ and $y \in F_{k+1}^c$, we have $\operatorname{dist}(x, E) > 2^{-k}$ and $\operatorname{dist}(y, E) \leq 2^{-k-1}$. It follows that $d(x, y) \geq 2^{-k-1}$ and thus for all $k \in \mathbb{N}$,

$$\inf_{x \in F_k} \operatorname{dist}(x, F_{k+1}^c) \geq 2^{-k-1} > 0.$$

By the previous homework, this implies that

$$\mu^*(A \cap E^c) = \mu^*\left(\bigcup_{k=0}^{\infty}(A \cap F_k)\right) = \lim_{k \to \infty} \mu^*(A \cap F_k).$$

In the mean time, for each $k \in \mathbb{N}$ we have $\inf_{x \in A \cap F_k} \operatorname{dist}(x, A \cap E) > 0$. This implies that

$$\mu^*(A \cap F_k) = \mu^*(A \cap E) + \mu^*(A \cap F_k).$$

Therefore, for each $k \in \mathbb{N}$, we have

$$\begin{aligned} \mu^*(A) &\geq \mu^*((A \cap E) \cup (A \cap F_k)) \\ &= \mu^*(A \cap E) + \mu^*(A \cap F_k). \end{aligned}$$

Taking the limit as $k \to \infty$ gives

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

This shows that E is measurable. Since all closed sets are in $\mathfrak{M}$ and $\mathcal{B}_X$ is the smallest σ-algebra containing all closed sets, we have $\mathcal{B}_X \subset \mathfrak{M}$.

($\impliedby$). Suppose $\mathcal{B}_X \subset \mathfrak{M}$. Let $A, B \subset X$ be such that $\Delta = \inf\{d(a, b) : a \in A, b \in B\} > 0$. Now define $E \subset X$ by

$$E = \left\{x \in X : \operatorname{dist}(x, A) < \frac{\Delta}{2}\right\}.$$

This is an open set, so $E \in \mathfrak{M}$. Additionally, $A \subset E$ and $B \subset E^c$. It follows that

$$\begin{aligned} \mu^*(A \cup B) &= \mu^*(A \cup B \cap E) + \mu^*(A \cup B \cap E^c) \\ &= \mu^*(A) + \mu^*(B). \end{aligned}$$

This implies that μ^* is an metric outer measure. □

Theorem. Let (X, d) be a metric space with gauge $(\mathcal{E}, \gamma)$ and outer measures $\mu_\delta^* : \mathcal{P}(X) \to [0, \infty]$ produced by $(\mathcal{E}_\delta, \gamma_\delta)$ for $\delta > 0$. Define $\mu_d^* : \mathcal{P}(X) \to [0, \infty]$ by

$$\mu_d^*(A) = \sup_{\delta > 0} \mu_d^*(A).$$

Then μ_d^* is a metric outer measure. Moreover, $\mu_d^*(A) = \lim_{\delta \to 0} \mu_\delta^*(A)$ for $A \subset X$.

Proof. ***** TO-DO ***** □

Definition. We call μ_d^* the metric outer measure generated by $(\mathcal{E}, \gamma)$.

Lemma. Let X be a set with gauge $(\mathcal{E}, \gamma)$ that covers X. Let $A \subset X$, then the following holds:

1. Let μ^* be the outer measure generated by $(\mathcal{E}, \gamma)$. Then there exists collection $\{E_{m,n}\}_{m,n=0}^{\infty} \subset \mathcal{E}$ such that $E = \bigcap_{m=0}^{\infty} \bigcup_{n=0}^{\infty} E_{m,n}$ such that $A \subset E$ and $\mu^*(A) = \mu^*(E)$.
2. Suppose (X, d) is metric space and the gauge is fine. Let μ_d^* be the metric outer measure. Then there exists collection $\{E_{m,n}\}_{m,n=0}^{\infty} \subset \mathcal{E}$ such that $E = \bigcap_{m=0}^{\infty} \bigcup_{n=0}^{\infty} E_{m,n}$ such that $A \subset E$ and $\mu^*(A) = \mu^*(E)$.

Proof. The proof for (1) is very similar to the proof for (2), so we only show (2) as follows. Since the gauge is fine, $(\mathcal{E}_\delta, \gamma_\delta)$ covers X for all $\delta > 0$. Then, for any $m \in \mathbb{N}$, there exists $\{E_{m,n}\}_n \subset \mathcal{E}_{2^{-m}}$ such that $A \subset \bigcup_{n=0}^{\infty} E_{m,n}$ and $\sum_{n=0}^{\infty} \gamma(E_{m,n}) \leq \mu_{2^{-m}}^*(A) + 2^{-m}$. Now let $E = \bigcap_{m=0}^{\infty} \bigcup_{n=0}^{\infty} E_{m,n}$. Note that $A \subset E$ and for any $m \in \mathbb{N}$, we have

$$\mu_{2^{-m}}^*(E) \leq \mu_{2^{-m}}^*\left(\bigcup_{n=0}^{\infty} E_{m,n}\right) \leq \sum_{n=0}^{\infty} \gamma(E_{m,n}) \leq \mu_{2^{-m}}^*(A) + 2^{-m}.$$

Taking the limit as $m \to \infty$, we have

$$\mu_d^*(E) \leq \mu_d^*(A) \leq \mu_d^*(E),$$

as desired.

□

Theorem. Let (X, d) be metric space with $(\mathcal{E}, \gamma)$ such that all sets in $\mathcal{E}$ are open. Assume that μ^* is a metric outer measure on X such that either

1. μ^* is generated by $(\mathcal{E}, \gamma)$, or
2. $\mu^* = \mu_d^*$ is generated by $(\mathcal{E}_\delta, \gamma_\delta)$.

Further suppose that $X = \bigcup_{n=0}^{\infty} A_n$ where $A_n \subset X$ is such that $\mu^*(A_n) < \infty$. Then the following holds:

1. The gauge covers X in case 1 and is fine in case 2.
2. In both cases, μ^* is cover-regular. More precisely, for each $A \subset X$, there is $G \in G_\delta(X) \subset \mathfrak{B}(X) \subset \mathfrak{M}$ such that $A \subset G$ and $\mu^*(A) = \mu^*(G)$.
3. In both cases, the following are equivalent for $E \subset X$:
   (a) $E \in \mathfrak{M}$, i.e. E is measurable.
   (b) there exists $G \in G_\delta(X)$ such that $E \subset G$ and $\mu^*(G \setminus E) = 0$.

(c) there exists $F \in F_\sigma(X)$ such that $F \subset E$ and $\mu^*(E \setminus F) = 0$.

Proof. **Step 0: proof for (1) and (2).**

We know $X = \bigcup_{n=0}^{\infty} A_n$ for some $\mu^*(A_n) < \infty$. For case (1), we can pick $\{E_{n,m}\} \subset \mathcal{E}$ such that $A_n \subset \bigcup_{m=0}^{\infty} E_{n,m}$. Then $X = \bigcup_{n=0}^{\infty} A_n = \bigcup_{n,m} E_{n,m}$. Therefore, $\mathcal{E}$ covers X. For case (2), note that $\mu_d^*(A_n) < \infty$ and $\mu_d^*(A_n) \geq \mu_\delta^*(A_n)$ for each $\delta > 0$ and $n \in \mathbb{N}$. Then for each $\delta > 0$, there exists $\{E_{n,m}\} \subset \mathcal{E}_\delta$ such that $A_n \subset \bigcup_{m=0}^{\infty} E_{n,m}$. It follows that $X = \bigcup_{n=0}^{\infty} A_n = \bigcup_{n,m} E_{n,m}$. Therefore, $(\mathcal{E}, \gamma)$ is fine.

We have the following observations:

1. μ^* is a metric outer measure. This implies that $\mathfrak{B}(X) \subset \mathfrak{M}$.
2. $G_\delta(X) \cup F_\sigma(X) \subset \mathfrak{B}(X) \subset \mathfrak{M}$ and $\mu^*(A) = 0$ implies $A \in \mathfrak{M}$.
3. By previous lemma and all sets in $\mathcal{E}$ are open, we know for each $A \subset X$ there is $E \in G_\delta(X)$ such that $A \subset E$ and $\mu^*(A) = \mu^*(E)$. In particular, μ^* is cover regular.

Step 1: starting on (3).

For (b) $\implies$ (a), suppose (b) holds for $E \subset X$. Then $E = G \setminus (G \setminus E) \in \mathfrak{M}$ since $\mu^*(G \setminus E) = 0$.

For (c) $\implies$ (a), suppose (c) holds for $E \subset X$. Then $E = F \cup (E \setminus F) \in \mathfrak{M}$ since $\mu^*(E \setminus F) = 0$.

Next we show "(a) $\implies$ (c)" implies "(a) $\implies$ (b)". Suppose $E \in \mathfrak{M}$, then $E^c \in \mathfrak{M}$. By (a) $\implies$ (b) we know there exists $F \in F_\sigma$ such that $F \subset E^c$ and $\mu^*(E^c \setminus F) = 0$. Let $G = F^c \in G_\delta$ then $E \subset G$ and $G \subset E = E^c \subset F$.

Therefore, it remains to show (a) $\implies$ (c) to complete the proof for the theorem.

Step 2: reduction for (a) $\implies$ (c).

Claim it suffices to show it for E such that $\mu^*(E) < \infty$. Suppose we did this and $\mu^*(E) = \infty$. Using observation there exists $B_n \in \mathfrak{M}$ such that $A_n \subset B_n$ and $\mu^*(B_n) = \mu^*(A_n) < \infty$. Then $E_n = E \cap B_n \in \mathfrak{M}$ and $\mu^*(E_n) < \infty$. Then by special case there is $F_n \in F_\sigma(X)$ such that $F_n \subset E_n$ and $\mu^*(F_n \setminus E_n) = 0$. Let $F = \bigcup_{n=0}^{\infty} F_n \in F_\sigma$ then $F \subset \bigcup_{n=0}^{\infty} E_n = E$ and

$$\mu^*(E \setminus F) \leq \sum_{n=0}^{\infty} \mu^*(E_n \setminus F_n) = 0.$$

Step 3: further reduction.

Claim it suffices to show it for the case where $\mu^*(E) < \infty$ and $E \in G_\delta(X)$. Suppose we have proved this and consider $E \subset X$ such that $\mu^*(E) < \infty$. Observation 3 allows us to pick $G \in G_\delta(X)$ such that $E \subset G$ and $\mu^*(E) = \mu^*(G)$. Now pick $H \in G_\delta$ such that $G \setminus E \subset H$ and $\mu^*(H) = \mu^*(G \setminus E)$.

Now apply special case. This gives $F \in F_\sigma$ such that $F \subset G$ and $\mu^*(G \setminus F) = 0$. Let $K = F \setminus H = F \cap H^c \in F_\sigma$ and $K = F \cap H^c \subset G \cap (G \setminus E)^c \subset E$.

Note that $E, F, G, H, K \in \mathfrak{M}$, so

$$\begin{aligned}\mu^*(E \setminus K) &= \mu^*(E) - \mu^*(K) \\ &= \mu^*(G) - \mu^*(F \setminus H) \\ &= \mu^*(G) - \mu^*(F) + \mu^*(F \cap H) \\ &\leq \mu^*(G) - \mu^*(F) + \mu^*(H) \\ &= \mu^*(G \setminus F) + \mu^*(H) \\ &= \mu^*(G \setminus E) \\ &= \mu^*(G) - \mu^*(E) \\ &= 0.\end{aligned}$$

Therefore, K is the desired F_σ set.

Step 4: finishing (a) $\implies$ (c).

Suppose $E \in G_\delta(X)$ and $\mu^*(E) < \infty$. Write $E = \bigcup_{n=0}^\infty V_n$ where $V_n \subset X$ open. For $m, n \in \mathbb{N}$, let

$$C_{n,m} = \left\{x \in V_n : \operatorname{dist}(x, V_n^c) \geq 2^{-m}\right\} \subset V_n.$$

Note that $C_{n,m}$ is closed, $C_{n,m} \subset C_{n,m+1}$, $V_n = \bigcup_m C_{n,m}$. Since $E, C_{n,m}, V_n \in \mathfrak{M}$, we have

$$\mu^*(E) = \mu^*(E \cap V_n) = \lim_{m \to \infty} \mu^*(E \cap C_{n,m}).$$

Thus, there exists $M(n,k)$ such that $\mu^*(E \setminus C_{n,M(n,k)}) < 2^{-n-k}$. Now let $D_k = \bigcup_{n=0}^\infty C_{n,M(n,k)}$ closed. Also, $D_k \subset \bigcup_{n=0}^\infty V_n = E$ and

$$\mu^*(E) - \mu^*(D_k) = \mu^*(E \setminus D_k) \leq \sum_{n=0}^\infty \mu^*(E \setminus C_{n,M(n,k)}) \leq 2^{-k+1}.$$

Let $F = \bigcup_{k=0}^\infty D_k \subset E$ and note that $F \in F_\sigma$. Then

$$\mu^*(E \setminus F) = \mu^*(E) - \mu^*(F) \leq \mu^*(E) - \mu^*(D_k) < 2^{-k+1}$$

for all $k \in \mathbb{N}$. Therefore, $\mu^*(E \setminus F) = 0$.

□

Lemma. Suppose (X, d) metric space with metric outer measure μ^*. Suppose $X = \bigcup_{n=0}^\infty V_n$ for $V_n \subset X$ open and $\mu^*(V_n) < \infty$. Suppose $E \subset G \in G_\delta(X)$ such that $\mu^*(G \setminus E) = 0$. Then for each $\varepsilon > 0$, there exists open $U \subset X$ such that $E \subset U$ and $\mu^*(U \setminus E) < \varepsilon$.

Proof. Let $E_n = E \cap V_n$ and $G = G \cap V_n$. Write $G = \bigcap_{j=0}^\infty W_j$ where W_j open. Now set

$$Z_{n,m} = V_n \cap \bigcap_{j=0}^m W_j,$$

which are open for all $n, m \in \mathbb{N}$. Now notice that $G_n \subset Z_{n,m+1} \subset Z_{n,m} \subset V_n$. Note that $\mu^*(V_n) < \infty$, so $\mu^*(G_n) = \lim_{m\to\infty} \mu^*(Z_{n,m})$. Therefore, for all $\varepsilon > 0$, there exists $M(n)$ such that

$$\mu^*(Z_{n,M(n)} \setminus G_n) < \varepsilon 2^{-n-2}.$$

Then set $U = \bigcup_{n=0}^{\infty} Z_{n,M(n)} \supset \bigcup_{n=0}^{\infty} G_n = G \supset E$ open, then we have

$$\begin{aligned}
\mu^*(U \setminus E) &= \mu^*(U \setminus G) + \mu^*(G \setminus E) \\
&= \mu^* \left(\bigcup_{n=0}^{\infty} Z_{n,M(n)} \cap \bigcap_{n=0}^{\infty} G_n^c \right) \\
&\le \sum_{n=0}^{\infty} \mu^*(Z_{n,M(n)} \setminus G_n) \\
&< \varepsilon,
\end{aligned}$$

as desired.

□

Definition (Outer-regular)**.** Let X be a metric space, $\mathfrak{M}$ a σ-algebra with $\mathfrak{B}(X) \subset \mathfrak{M}$ and suppose $\mu : \mathfrak{M} \to [0, \infty]$ is a measure. Say μ is outer-regular if

$$\mu(E) = \inf \{\mu(U) : E \subset U \text{ open}\}.$$

3.2 Lebesgue and Hausdorff measure

***** TO-DO *****

3.3 Measurable and μ-measurable functions

Definition (Measurable functions)**.** Let $(X, \mathfrak{M})$ and $(Y, \mathfrak{N})$ be measurable sets. A map $f : X \to Y$ is called $(\mathfrak{M}, \mathfrak{N})$ measurable if $f^{-1}(E) \in \mathfrak{M}$ for all $E \in \mathfrak{N}$.

***** TO-DO *****

Definition (Simple functions)**.** Let $(X, \mathfrak{M})$ and $(Y, \mathfrak{N})$ be measurable sets. A map $f : X \to Y$ is called simple if it is measurable and $f(X)$ is finite. Write the set of all simple functions from X to Y as $S(X, Y)$.

Theorem (Characterization of $\overline{\mathbb{R}}$ measurablility)**.** Let $(X, \mathfrak{M})$ be measure space and $f : X \to \overline{\mathbb{R}}$. The following are equivalent:

1. f is measurable.
2. There exists $\{\phi_k\}_{k=0}^{\infty} \subset S(X; \overline{\mathbb{R}})$ such that $\phi_k \to f$ pointwise as $k \to \infty$.

Moreover, if f is measurable, the sequence can be built such that

– On the set $\{f \geq 0\}$, we have $0 \leq \phi_k \leq \phi_{k+1} \leq f$.

– On the set $\{f < 0\}$, we have $f \leq \phi_{k+1} \leq \phi_k \leq 0$.

– If f is actually from X to $\mathbb{R}$ and is bounded, then $\phi_k \to f$ uniformly.

Proof. (2) $\implies$ (1). Pointwise limit of measurable functions are measurable.

(1) $\implies$ (2). Suppose $f : X \to [0, \infty]$ is measurable. For $k \in \mathbb{N}$, define $\phi_k : [0, \infty)$ by

$$\phi_k(x) = \begin{cases} (j-1)2^{-k} & \text{if } (j-1)2^{-k} \leq f(x) < j2^{-k} \text{ for } 1 \leq j \leq k2^k, \\ k & \text{if } f(x) > k. \end{cases}$$

Because f is measurable, ϕ_k is simple for each $k \in \mathbb{N}$.

Note that $0 \leq \phi_k \leq \phi_{k+1} \leq f$. Also, if $f(x) < \infty$, then $0 \leq f(x) - \phi_k(x) \leq 2^{-k}$. If $f(x) = \infty$, then $\phi_k(x) = k$. This shows that $\phi_k \to f$. Moreover, if f is bounded then $\phi_k \to f$ uniformly.

In the general case, apply the special case to f on $\{f \geq 0\}$ and $-f$ on $\{f < 0\}$.

□

Definition (Separably-valued)**.** Let X be a set and Y a metric space. A map $f : X \to Y$ is **separably-valued** if $f(X) \subset Y$ is separable.

Theorem. Let $(X, \mathfrak{M})$ be measure space and Y be metric space, $f : X \to Y$. The following are equivalent for $f : X \to Y$:

1. f is $(\mathfrak{M}, \mathfrak{B}(Y))$ measurable and separably valued.
2. There exists $\{\phi_k\}_{k=0}^{\infty} \in S(X; Y)$ such that $\phi_k \to f$ pointwise.

Proof. (2) $\implies$ (1). The pointwise limit of measurable function is measurable. On the other hand, $f(X) = \overline{\bigcup_{k=0}^{\infty} \phi_k(X)}$, which is separable since $\phi_k(X)$ finite for any $k \in \mathbb{N}$.

(1) $\implies$ (2). Assume initially that Y is totally bounded. Then for each $n \in \mathbb{N}$ there exists $y_0^n, \ldots, y_{K(n)}^n \in Y$ such that $Y = \bigcup_{k=0}^{K(n)} B(y_k^n, 2^{-n})$. Let $V_0^n = B(y_0^n, 2^{-n})$ and for $k \geq 1$ define $V_k^n = B(y_k^n, 2^{-n}) \setminus \bigcup_{j=0}^{k-1} B(y_j^n, 2^{-n})$. Then, $Y = \bigsqcup_{k=0}^{M(n)} V_k^n$ where $V_k^n = \emptyset$ for $M(n) < k \leq K(n)$.

Define $\phi_n : Y \to \{y_0^n, \ldots, y_{M(n)}^n\}$ via $\phi_n(y) = y_k^n$ if $y \in V_k^n$. Clearly ϕ_n is simple and $d(\phi_n(y), y) < 2^{-n}$ for all $n \in \mathbb{N}$ and $y \in Y$. Therefore, $\phi_n(y) \to (y)$ pointwise. Then $f_n = \phi_n \circ f$ are simple functions from X to Y. Also, since $\phi_n \to \text{id}$ pointwise, $f_n \to f$ pointwise.

Now consider the general case in which $f(X)$ is a separable subset of Y. Then there exists a homeomorphism $h : f(X) \to Z$ for Z a totally bounded metric space, for example take Z a subset of Hilbert cube H^∞ since all separable metric space is homeomorphism to a subset of the Hilbert cube. Thus $h \circ f : X \to Z$ is measurable with Z totally bounded, so the special case provides a sequence $\{\phi_n\}_{n=0}^{\infty} \subset S(X; Z)$ such that $\phi_n \to h \circ f$ pointwise. Then, $h^{-1} \circ \phi_n \in S(X; Y)$ is such that $h^{-1} \circ \phi_n \to h^{-1} \circ h \circ f = f$ pointwise, using continuity of h and h^{-1}.

□

Definition (Almost everywhere)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $P(x)$ be a proposition for every $x \in X$. Say P is true **almost everywhere** (a.e.) if there exists a set $N \in \mathfrak{M}$ such that $\mu(N) = 0$ and $P(x)$ is true for all $x \in N^c$.

Theorem. Let $(X, \mathfrak{M}, \mu)$ be a measure space. Let Y be a metric space, $f : X \to Y$. The following are equivalent:

1. There exists $\{\psi_n\}_{n=0}^{\infty} \subset S(X; Y)$ such that $\psi_n \to f$ pointwise a.e. in X.
2. There exists a measurable and separably valued $F : X \to Y$ such that $f = F$ a.e.
3. There exists a null set $N \in \mathfrak{M}$ and a measurable $F : X \to Y$ such that $f = F$ on N^c and $f(N^c)$ is separable in Y.

Proof. (1) $\implies$ (2). There exists $N \in \mathfrak{M}$ null such that $\psi_n \to f$ pointwise in N^c. Thus, $f : N^c \to Y$ is measurable and separably valued by the previous theorem. Note the constant map $N \ni x \mapsto y \in Y$ for $y \in Y$ fixed is measurable. Thus we can define $F : X \to Y$ by

$$F(x) = \begin{cases} f(x) & (x \in N^c), \\ y & (x \in N). \end{cases}$$

Then F is measurable. It is also separably valued since $F(X) = f(N^c) \cup \{y\}$.

(2) $\implies$ (3). Trivial.

(3) $\implies$ (1). Note that $F : N^c \to Y$ is measurable and $F(N^c) = f(N^c)$ is separable. By previous theorem, there exists $\{\phi_n\}_{n=0}^{\infty} \in S(N^c; Y)$ such that $\phi_n \to F = f$ pointwise on N^c. Now let $\psi_n \in S(X; Y)$ be ϕ_n in N^c and $y \in Y$ fixed in N. Then $\psi_n \to f$ pointwise in N^c.

□

Definition. Let $(X, \mathfrak{M})$ be measurable, Y be either a normed vector space or $\overline{\mathbb{R}}$. Let $\psi \in S(X; Y)$.

1. A **representation** of ψ is a finite and well-defined sum $\psi = \sum_{k=1}^{K} v_k \chi_{E_k}$ for $v_k \in Y$ and $E_k \in \mathfrak{M}$.
2. A **canonical representation** is $\psi = \sum_{v \in \psi(X)} v \chi_{\psi^{-1}(\{v\})}$
3. Now suppose μ is a measure. We say a representation $\psi = \sum_{k=1}^{K} v_k \chi_{E_k}$ is **finite** if $\mu(E_k) < \infty$ for all k such that $v_k \neq 0$. We say ψ is a **finite simple function** if it has a finite representation.

   We write $S_{\text{fin}}(X; Y) = \{f \in S(X; Y) : f \text{ is finite}\}$. Note that it is clear ψ is finite if and only if the canonical representation is finite if and only if $\mu(\text{spt}(\psi)) < \infty$ where $\text{spt}(\psi) = \{x \in X : \psi(x) \neq 0\}$ is the set-theoretic support of ψ.

Definition. Let $(X, \mathfrak{M}, \mu)$ be a measure space and Y be a metric space.

1. We say $f : X \to Y$ is **almost measurable** if $f = F$ a.e. with $F : X \to Y$ is measurable.
2. We say $f : X \to Y$ is **almost separably valued** if there exists a null set $N \in \mathfrak{M}$ such that $f(N^c)$ is separable.
3. We say $f : X \to Y$ is **μ-measurable** if it is almost measurable and almost separably valued. Equivalently, f is the a.e. limit of simple functions.
4. Suppose Y is a normed vector space or $\overline{\mathbb{R}}$. We say $f : X \to Y$ is **strongly μ-measurable** if there exists $\{\psi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; Y)$ such that $\psi_n \to f$ a.e. as $n \to \infty$.

Example. Let $X = \{1, 2, 3\}$ and $\mathfrak{M} = \{\emptyset, \{1, 2\}, \{3\}, \{1, 2, 3\}\}$. Let $f, g : X \to \mathbb{R}$ via $f(x) = x$ and $g(x) = 3$. Then f is not measure since $f^{-1}(\{1\}) = \{1\} \notin \mathfrak{M}$ but g is measurable.

Now equip $(X, \mathfrak{M})$ with the measure δ_3. Then, $f = g$ a.e. This shows that equality almost everywhere does not preserve measurablility. The problem is that $(X, \mathfrak{M}, \delta_3)$ is not **complete**. △

This brings us to the next theorem.

Theorem. Let $(X, \mathfrak{M}, \mu)$ be a measure space. Then the following are equivalent:

1. $(X, \mathfrak{M}, \mu)$ is complete.
2. If $(Y, \mathfrak{N})$ is a measure space, $f, g : X \to Y$, f is measurable and $f = g$ a.e. then g is measurable.
3. If Y is a metric space with $\operatorname{card} Y = 2$, $f, g : X \to Y$, f measurable, $f = g$ a.e. then g is measurable.

Proof. (1) $\implies$ (2). Suppose $f, g : X \to Y$, f is measurable, $f = g$ a.e. Pick null set $N \in \mathfrak{M}$ such that $f = g$ on N^c. Take $E \in \mathfrak{N}$, then

$$\begin{aligned} g^{-1}(E) &= (g^{-1}(E) \cap N) \cup (g^{-1}(E) \cap N^c) \\ &= (g^{-1}(E) \cap N) \cup (f^{-1}(E) \cap N^c). \end{aligned}$$

Note that $f^{-1}(E) \cap N^c$ is measurable, and $g^{-1}(E) \cap N \subset N$ null, so it is also measurable. Therefore, $g^{-1}(E)$ is measurable and g is measurable.

(2) $\implies$ (3). Clear.

(3) $\implies$ (1). Prove the contrapositive. Suppse $(X, \mathfrak{M}, \mu)$ is not complete and $Y = \{y, z\}$ a metric space. Find $\emptyset \neq A \subsetneq B$ such that $\mu(B) = 0$ and $A \notin \mathfrak{M}$. Define $f, g : X \to Y$ by

$$g(x) = \begin{cases} y & (x \notin A), \\ z & (x \in A). \end{cases}$$

and $f(x) = y$ be constant. Then $f = g$ a.e. f is measurable, and g is not measurable.

□

Corollary. Let $(X, \mathfrak{M}, \mu)$ be a complete measurable space, Y a separable metric space, and $f : X \to Y$. Then, f is μ-measurable if and only if f is measurable.

Proposition. Let $(X, \mathfrak{M}, \mu)$ be a measure space and Y be a metric space. The following holds:

1. Let $f, g : X \to Y$. If f is μ-measurable and $f = g$ a.e. then g is μ-measurable.
2. Suppose Y is a normed vector space or $\overline{\mathbb{R}}$. If $f, g : X \to Y$, f is strongly μ-measurable, $f = g$ a.e. then g is strong μ-measurable.

Proof.

1. Let $\{\phi_n\}_{n=0}^{\infty} \subset S(X; Y)$ be such that $\phi_n \to g$ pointwise a.e. Pick null set $N \in \mathfrak{M}$ such that $f = g$ on N^c. Pick null set $Z \in \mathfrak{M}$ such that $f = \lim_{n\to\infty} \phi_n$. This implies that $g = \lim_{n\to\infty} \phi_n$ on $(N \cup Z)^c$.
2. Same proof as the first item but let $\{\phi_n\}_{n=0}^{\infty} \in S_{\text{fin}}(X; Y)$.

□

Theorem. Let $(X, \mathfrak{M}, \mu)$ be a measure space and Y be a normed vector space with $V \neq \{0\}$. Then the following are equivalent:

1. $(X, \mathfrak{M}, \mu)$ is σ-finite.
2. If $f : X \to Y$ is μ-measurable, then f is strongly μ-measurable.
3. Let $f : X \to Y$, then f is μ-measurable if and only if f is strongly μ-measurable.
4. If $y \in Y \setminus \{0\}$, then $f : X \to Y$ via $f(x) = y$ strongly μ-measurable.

Proof. (1) $\implies$ (2). Suppose $(X, \mathfrak{M}, \mu)$ is σ-finite. We can find $\{X_n\}_{n=0}^{\infty} \subset \mathfrak{M}$ such that $X_n \subset X_{n+1}$, $\mu(X_n) < \infty$ and $\bigcup_{n=0}^{\infty} X_n = X$. Let $f : X \to Y$ be μ-measurable. Pick $\{\psi_n\}_{n=0}^{\infty} \subset S(X; Y)$ such that $\psi_n \to f$ pointwise a.e. Define $\phi_n = \chi_{X_n} \psi_n$. This shows that f is strongly μ-measurable.

(2) $\iff$ (3). Trivial since strongly μ-measurablility implies μ-measurablility.

(2) $\implies$ (4). Constant function are μ-measurable.

(4) $\implies$ (1). Let $y \in Y \setminus \{0\}$ and define $f : X \to Y$ via $f(x) = y$. This is strongly μ-measurable by assumption. Then there exists $\{\phi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; Y)$ such that $\phi_n \to f$ pointwise on N^c where N is null.

Pick $\varepsilon > 0$ such that $\{0\} \cap B(y, \varepsilon) = \emptyset$. Set $X_n = \phi_n^{-1}(B(y, \varepsilon))$. Then we have $\mu(X_n) < \infty$. For any $x \in N^c$ and n sufficiently large, $\phi_n(x) \in B(y, \varepsilon)$. Therefore, $N^c \subset \bigcup_{n=0}^{\infty} X_n$ and the proof we are complete.

□

Finally, we present a useful characterization of μ-measurablility of Banach-valued maps.

Theorem (Pettis)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space and V be a Banach space over $\mathbb{F}$. Suppose $W \subset V^*$ is a norming subspace. Let $f : X \to V$. Then the following are equivalent:

1. f is μ-measurable.
2. f is almost separably valued, and $w \circ f : X \to \mathbb{F}$ is μ-measurable for each $w \in V^*$.
3. f is almost separably valued, and $w \circ f : X \to \mathbb{F}$ is μ-measurable for each $w \in W$.

In any case, there exists $\{\phi_n\}_{n=0}^\infty \subset S(X; V)$ such that $\|\phi_n\| \leq 2\|f\|$ on X such that $\phi_n \to f$ pointwise a.e. as $n \to \infty$. Moreover, the same equivalence holds with μ-measurablility replaced by strongly μ-measurablility and $\{\phi_n\}_{n=0}^\infty$ replaced by $\{\phi_n\}_{n=0}^\infty \subset S_{\text{fin}}(X; V)$.

Proof. (1) $\implies$ (2). Suppose f is μ-measurable, which means it is almost separably valued. Each $w \in V^*$ is also continuous so $w \circ f$ is μ-measurable.

(2) $\implies$ (3). Trivial since $W \subset V^*$.

(3) $\implies$ (1). Suppose f is almost separably valued. Then there exists null set $N_* \subset X$ such that $f(X \setminus N_*) \subset V$ separable. Define the subspace

$$M = \operatorname{span}(f(X \setminus N_*)) \subset V,$$

which is separable by construction. Pick a dense set $\{v_n\}_{m=0}^\infty \subset M$ such that $v_0 = 0$. Then by a previous theorem, we know there exists a norming sequence $\{w_n\}_{n=0}^\infty \subset W$ for M.

Now, given any $v \in V$ and $n \in \mathbb{N}$, define the function $\Phi_{n,v} : X \to [0, \infty)$ by

$$\Phi_{n,v}(x) = |\langle w_n, f(x) - v\rangle| = |w_n(f(x) - v)|.$$

Note that $X \ni x \mapsto \langle w_n, v\rangle \in \mathbb{F}$ is μ-measurable and the map $X \ni x \mapsto \langle w_n, f(x)\rangle \in \mathbb{F}$ is also μ-measurable by assumption. It follows that $\Phi_{n,v}$ is μ-measurable. Therefore, there exists null set $N_{n,v} \subset X$ and a measurable map $\Psi_{n,v} : X \to [0, \infty)$ such that $\Psi_{n,v} = \Phi_{n,v}$ on $X \setminus N_{n,v}$. For each $v \in V$ define null set

$$N(v) = N_* \cup \bigcup_{n=0}^\infty N_{n,v} \subset X,$$

with $\Psi_{n,v} = \Phi_{n,v}$ on $X \setminus N(v)$ for all $n \in \mathbb{N}$.

For $v \in M$ define the map $\Phi_v : X \to [0, \infty]$ by $\Phi_v(x) = \|f(x) - v\|$ and note that $\{w_n\}_{n=0}^\infty$ is norming sequence for M. This implies that

$$\Phi_v(x) = \sup_{n \in \mathbb{N}} |\langle w_n, f(x) - v\rangle|$$

for all $x \in X \setminus N_*$. We also have that

$$\Phi_v(x) = \sup_{n \in \mathbb{N}} \Phi_{n,v}(x) = \sup_{n \in \mathbb{N}} \Psi_{n,v}(x)$$

for all $x \in X \setminus N(v)$, so Φ_v is measurable when restricted to $X \setminus N(v)$. We can then define the set

$$N = \bigcup_{m=0}^{\infty} N(v_m) \subset X,$$

which is null. By construction, each Φ_{v_m} is measurable when restricted to N^c. In particular, $\Phi_0 = \Phi_{v_0} = \|f\|$ is measurable when restricted to N^c.

For $u \in M$ and $n \in \mathbb{N}$, define

$$k(n,u) = \min\left\{0 \leq k \leq n : \|u - v_k\| = \min_{0 \leq j \leq n} \|u - v_j\|\right\}.$$

By construction,

$$\left\|v_{k(n,u)}\right\| \leq \left\|u - v_{k(n,m)}\right\| + \|u\| \leq \|u - v_0\| + \|u\| = 2\|u\|.$$

We then define $S_n : M \to \{v_0, \ldots, v_n\}$ via $S_n(u) = v_{k(n,u)}$. Note that $\|S_n(u)\| \leq 2\|u\|$. Also, $\{v_m\}_{m=0}^{\infty}$ dense in M implies $S_n(u) \to u$ as $n \to \infty$.

Finally, for $n \in \mathbb{N}$, define $\psi_n : N^c \to \{v_0, \ldots, v_n\} \subset V$ via $\psi_n = S_n \circ f$. For $0 \leq k \leq n$, we compute

$$\begin{aligned}
&\{x \in N^c : \psi_n(x) = v_k\} \\
&= \left\{x \in N^c : \|f(x) - v_k\| = \min_j \|f(x) - v_j\|\right\} \cap \bigcap_{j=0}^{k-1} \{x \in N^c : \|f(x) - v_k\| < \|f(x) - v_j\|\}
\end{aligned}$$

This set is measurable since Φ_{v_m} measurable on N^c for each $m \in \mathbb{N}$. It follows that ψ_n is measurable on N^c. Let $\phi_n \in S(X;V)$ by

$$\phi_n(x) = \begin{cases} \psi_n(x) & (x \in N^c), \\ 0 & (x \in N). \end{cases}$$

Then, $\|\phi_n\| \leq 2\|f\|$ and $\phi_n(x) = \psi_n(x) \to f(x)$ as $n \to \infty$ for $x \in N^c$. Therefore, $\phi_n \to f$ a.e. and thus f is μ-measurable.

□

3.4 Lebesgue-Bochner Integral

Lemma. Let $(X, \mathfrak{M}, \mu)$ be a measure space and $Y \in \{V, [0,\infty]\}$. Let $\psi : X \to Y$ be simple such that

$$\psi = \sum_{i=1}^{I} \alpha_i \chi_{E_i} = \sum_{j=1}^{J} \beta_j \chi_{F_j}.$$

Additionally, if $Y = V$ suppose both representation are finite. Then,

$$\sum_{i=1}^{I} \alpha_i \mu(E_i) = \sum_{j=1}^{J} \beta_j \mu(F_j).$$

Based on this lemma, we can define

$$\int_X \psi \, d\mu = \sum_{i=1}^{I} \alpha_i \mu(E_i).$$

This induces maps $\int_X \cdot \, d\mu : S(X; [0,\infty]) \to [0,\infty]$ and $\int_X \cdot \, d\mu : S_{\text{fin}}(X; V) \to V$.

Proposition. Let $(X, \mathfrak{M}, \mu)$ be a measure space and $Y \in \{V, [0,\infty]\}$. Then the following holds:

1. If $Y = V$, then
$$\int_X (\alpha f + \beta g) \, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu$$
for all $\alpha, \beta \in \mathbb{F}$ and $f, g \in S_{\text{fin}}(X; V)$. If $Y = [0,\infty]$, the same equality holds for any $\alpha, \beta > 0$ and $f, g \in S(X; V)$.
2. If $Y = V$, then $\|f\| \in S_{\text{fin}}(X; [0,\infty))$ and
$$\left\| \int_X f \, d\mu \right\| \leq \int_X \|f\| \, d\mu.$$
3. If $E \in \mathfrak{M}$, then
$$\int_E f \, d\mu = \int_X f \chi_E \, d\mu.$$
4. If $N \in \mathfrak{M}$ is a null set, then
$$\int_N f \, d\mu = 0.$$
5. If $A, B \in \mathfrak{M}$ is such that $A \cap B = \emptyset$, then
$$\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu.$$
6. Suppose $\{X_n\}_{n=0}^{\infty} \subset \mathfrak{M}$ is such that $X_n \subset X_{n+1}$ and $\mu(X_n) < \infty$. Then
$$\int_X f \, d\mu = \lim_{n\to\infty} \int_{X_n} f \, d\mu.$$

Proof. Write $f = \sum_k f_k \chi_{E_k}$ be the canonical representation. We then have

$$\int_{X_n} f\, d\mu = \sum_k f_k \mu(X_n \cap E_k).$$

For each k, we have $X_n \cap E_k \subset X_{n+1} \cap E_k$ and $\bigcup_{n=0}^{\infty}(X_n \cap E_k) = E_k$. It follows that

$$\lim_{n\to\infty} \mu(X_n \cap E_k) = \mu(E_k).$$

Therefore,

$$\lim_{n\to\infty} \int_{X_n} f\, d\mu = \sum_k f_k \mu(E_k) = \int_X f\, d\mu.$$

□

7. If $Y = \mathbb{R}$ or $Y = [0,\infty]$ and $f \le g$ a.e. then

$$\int_X f\, d\mu \le \int_X g\, d\mu.$$

3.4.1 Integration of real-valued functions

Note that if $(X, \mathfrak{M}, \mu)$ is a measure space and $\phi \in S(X; [0,\infty])$, then

$$\int_X \phi\, d\mu = \sup\left\{ \int_X \psi\, d\mu : \psi \in S(X;[0,\infty]) \text{ and } \psi \le \phi \text{ a.e.} \right\}.$$

Definition. Let $(X, \mathfrak{M}, \mu)$ be a measure space. Let $f : X \to [0,\infty]$ be μ-measurable. We define

$$\int_X f\, d\mu = \sup\left\{ \int_X \psi\, d\mu : \psi \in S(X;[0,\infty]) \text{ and } \psi \le f \text{ a.e.} \right\} \in [0,\infty].$$

We say f is **integrable** if $\int_X f\, d\mu < \infty$.

Remark. There are two remarks with regard to the definition above.

1. In principle we do not need f to be μ-measurable here. We build this into the definition because the resulting integral is more-or-less useless without this assumption.
2. $[0,\infty]$ is a separable metric space, so for $f : X \to [0,\infty]$, f is measurable implies f is μ-measurable, and f almost measurable implies f is μ-measurable.

Theorem. Let $(X, \mathfrak{M}, \mu)$ be a measure space, $f, g : X \to [0,\infty]$ be μ-measurable functions. The following holds:

1. For $\alpha \in [0,\infty)$, we have

$$\int_X \alpha f\, d\mu = \alpha \int_X f\, d\mu.$$

2. If $f \leq g$ a.e. then
$$\int_X f\,d\mu \leq \int_X g\,d\mu.$$
3. If $f = g$ a.e. then
$$\int_X f\,d\mu = \int_X g\,d\mu.$$
4. For $E \in \mathfrak{M}$, we have
$$\int_E f\,d\mu = \int_X f\chi_E\,d\mu.$$
5. If $N \in \mathfrak{M}$ is null, then
$$\int_N f\,d\mu = 0.$$

Proof. Follow directly from corresponding results in $S(X;[0,\infty])$ and the definition of $\int_X f\,d\mu$. $\square$

Theorem (Monotone convergence theorem, basic version)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space and suppose for each $n \in \mathbb{N}$, we have $f_n : X \to [0,\infty]$ **measurable**. Further suppose that $f_n \leq f_{n+1}$ on X and $f : X \to [0,\infty]$ is given by $f = \lim_{n\to\infty} f_n$. Then f is measurable and
$$\int_X f\,d\mu = \lim_{n\to\infty} \int_X f_n\,d\mu = \sup_{n\in\mathbb{N}} \int_X f_n\,d\mu.$$

Proof. We already know f is measurable. Also, $f_n \leq f_{n+1} \leq f$ on X, so
$$\int_X f_n\,d\mu \leq \int_X f_{n+1}\,d\mu \leq \int_X f\,d\mu.$$
It follows that
$$\lim_{n\to\infty} \int_X f_n\,d\mu \leq \int_X f\,d\mu.$$
To show the opposite inequality, let $\phi \in S(X;[0,\infty])$ such that $\phi \leq f$ a.e. and $\alpha \in (0,1)$. Let $N \in \mathfrak{M}$ be a null set and $\phi \leq f$ on N^c. Also, for each $n \in \mathbb{N}$, let $E_n = \{x \in X : f_n(x) \geq \alpha\phi(x)\}$. Note the following:

1. Since $f_n \leq f_{n+1}$, we have $E_n \subset E_{n+1}$.
2. Since $f_n \to f$ pointwise, we have $X = N \cup \bigcup_{n=0}^\infty E_n$.
3. We have
$$\alpha \int_{N\cup E_n} \phi\,d\mu = \int_{E_n} \alpha\phi\,d\mu \leq \int_{E_n} f_n\,d\mu \leq \int_X f_n\,d\mu$$
4. We have
$$\int_X \phi\,d\mu = \lim_{n\to\infty} \int_{N\cup E_n} \phi\,d\mu.$$

Therefore,

$$\alpha \int_X \phi \, d\mu = \lim_{n\to\infty} \alpha \int_{N\cup E_n} \phi \, d\mu \le \lim_{n\to\infty} \int_X f_n \, d\mu.$$

Since the above inequality holds for all $\alpha \in (0,1)$, we konw $\int_X \phi \, d\mu \le \lim_{n\to\infty} \int_X f_n \, d\mu$. This is then true for all simple function ϕ such that $\phi \le f$ a.e. Taking the sup gives

$$\int_X f \, d\mu \le \lim_{n\to\infty} f_n \, d\mu.$$

The proof is then complete. □

Theorem. Let $(X, \mathfrak{M}, \mu)$ be measure space, $f, g : X \to [0,\infty]$ be μ-measurable. Then

$$\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu.$$

Proof. Recall that μ-measurable functions are almost measurable. Choose measurable functions $F, G : X \to [0,\infty]$ such that $f = F$ and $g = G$ a.e. We may then choose $\{\phi_n\}_{n=0}^{\infty}, \{\psi_n\}_{n=0}^{\infty} \subset S(X;[0,\infty])$ such that $\lim_{n\to\infty} \phi_n = F$ and $\lim_{n\to\infty} \psi_n = G$, $0 \le \phi_n \le \phi_{n+1} \le F$ and $0 \le \psi_n \le \psi_{n+1} \le G$. Then

$$0 \le \phi_n + \psi_n \le \phi_{n+1} + \psi_{n+1} \le F + G = \lim_{n\to\infty} (\phi_n + \psi_n).$$

It follows then from monotone convergence theorem that

$$\begin{aligned}
\int_X (F+G) \, d\mu &= \lim_{n\to\infty} \int_X (\phi_n + \psi_n) \, d\mu \\
&= \lim_{n\to\infty} \int_X \phi_n \, d\mu + \lim_{n\to\infty} \int_X \psi_n \, d\mu \\
&= \int_X F \, d\mu + \int_X G \, d\mu.
\end{aligned}$$

Since $f = F$ and $g = G$ a.e. we have

$$\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu.$$

□

Recall: given $f : X \to \overline{\mathbb{R}}$, we write $f^{\pm} : X \to [0,\infty]$ via

$$f^+ = \max\{0, f\}, \quad f^- = \max\{0, -f\}.$$

Then we have $f = f^+ - f^-$ and $|f| = f^+ + f^-$. Also, if f is measurable or μ-measurable, then $f^{\pm}$ is also measurable or μ-measurable since they are composition of a continuous function (namely $x \mapsto \max\{0, x\}$) with a measurable or μ-measurable function.

Definition. Let $(X, \mathfrak{M}, \mu)$ be measure space and $f : X \to \overline{\mathbb{R}}$ be μ-measurable. If either f^+ or f^- is **integrable**, we say f is **extended integrable** and set

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu \in \overline{\mathbb{R}}.$$

We say f is **integrable** if $f^\pm$ are both integrable.

Proposition (absolute integrability)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space, $f : X \to \overline{\mathbb{R}}$ be μ-measurable. Then f is integrable if and only if $|f|$ is integrable.

Proof. We know f is integrable if and only if $f^\pm$ are both integrable, but $|f| = f^+ + f^-$. Therefore, f integrable implies $|f|$ is integrable. Conversely, if $|f|$ is integrable, then $0 \leq f^\pm \leq |f|$, so $f^\pm$ are both integrable. □

Theorem. Let $(X, \mathfrak{M}, \mu)$ be a measure space, $f, g : X \to \overline{\mathbb{R}}$ are extended integrable. The following holds:

1. For all $E \in \mathfrak{M}$, we have $\int_E f \, d\mu = \int_X f \chi_E \, d\mu$.
2. For all $\alpha \in \mathbb{R}$, we have $\alpha \int_X f \, d\mu = \int_X \alpha f \, d\mu$.
3. $\int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$, provided that all operations are well-defined.
4. $\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu$ for all $A, B \in \mathfrak{M}$ such that $A \cap B = \emptyset$.
5. If $f \leq g$ a.e. then $\int_X f \, d\mu \leq \int_X g \, d\mu$.
6. $\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu$.
7. If $|f| \leq g$ a.e. and g integrable, then f is integrable.

Theorem (Chebyshev's inequality)**.** If f is measurable and integrable, then

$$\mu\left(\{x \in X : |f(x)| \geq \alpha\}\right) \leq \frac{1}{\alpha} \int_X |f| \, d\mu$$

for all $\alpha \in (0, \infty)$.

Proof.

$$\text{LHS} = \int_{\{|f| \geq \alpha\}} 1 \, d\mu = \int_{\{|f| \geq \alpha\}} \frac{|f|}{\alpha} \, d\mu = \frac{1}{\alpha} \int_X |f| \, d\mu = \text{RHS}\,.$$

□

Corollary. Let $(X, \mathfrak{M}, \mu)$ be a measure space and $f : X \to \overline{\mathbb{R}}$.

1. If f is integrable, then there exists a null set $N \in \mathfrak{M}$ and a σ-finite set $E \in \mathfrak{M}$ such that $\{|f| = \infty\} \subset N$ and $\operatorname{spt}(f) \subset E$.

2. If f is extended integrable, then there exists a null set $N \in \mathfrak{M}$ such that either $\{f = \infty\} \subset N$ or $\{f = -\infty\} \subset N$.

Proof. 1. Suppose initially that f is measurable and integrable, then Chebyshev's inequality implies that

$$\mu\left(\{|f| = \infty\}\right) \leq \mu\left(\left\{|f| > 2^k\right\}\right) \leq 2^{-k} \int_X |f| \, d\mu$$

for all $k \in \mathbb{N}$. It follows that $\mu\left(\{|f| = \infty\}\right)$ is null.

On the other hand, $\operatorname{spt}(f) = \bigcup_{k=0}^{\infty} \left\{|f| > 2^{-k}\right\}$, but

$$\mu\left(\left\{|f| > 2^{-k}\right\}\right) \leq 2^k \int_X |f| \, d\mu < \infty.$$

It follows that $\operatorname{spt}(f)$ is σ-finite.

In general, if f is integrable and μ-measurable, pick $F = f$ a.e. for F measurable and integrable and apply the argument above.

2. Next, if f is extended integrable but not integrable, then either f^+ is integrable or f^- is integrable. If f^+ is integrable, then $\{f = +\infty\}$ is contained in some null set. If f^- is integrable, $\{f = -\infty\}$ is contained in a null set.

□

To prove the more general form of monotone convergence theorem, we first need a useful lemma.

Lemma. Let $(X, \mathfrak{M}, \mu)$ be a measure space and suppose that $f : X \to \overline{\mathbb{R}}$ is μ-measurable and $g : X \to \mathbb{R}$ is integrable. Further suppose $g \leq f$ a.e. Then, f and $f - g$ are extended integrable, and

$$\int_X (f - g) \, d\mu = \int_X f \, d\mu - \int_X g \, d\mu.$$

Proof. Since $g \leq f$ a.e. we have $f^- \leq g^-$ a.e. Since g is integrable, f^- is integrable and thus f is extended-integrable. We also have $f - g$ well defined on all of X and $f - g \geq 0$ a.e. Therefore, $f - g$ is extended-integrable.

If f is integrable, then we immediately have the desired equality. Suppose not f is not integrable but only extended-integrable. This implies f^+ is not integrable. We must then have $f - g$ not integrable, otherwise $f = (f - g) + g$ is integrable. Therefore, $\int_X (f - g) \, d\mu = \int_X f \, d\mu = \infty$, and the desired equality holds. □

Theorem (Monotone convergence theorem, general form)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space and suppose $f_k : X \to \overline{\mathbb{R}}$ is μ-measurable for all $k \in \mathbb{N}$. Suppose that $f : X \to \mathbb{R}$ is such that $f_k \to f$ a.e. Then, f is μ-measurable and the following holds:

1. Suppose that $\{f_k\}_{k=0}^{\infty}$ is almost everywhere nondecreasing, that is, $f_k \leq f_{k+1}$ a.e. Suppose also that there exists an integrable function $g : X \to \overline{\mathbb{R}}$ such that $g \leq f_k$ a.e. for all $k \in \mathbb{N}$. Then, f and f_k are extended integrable for all $k \in \mathbb{N}$, and

$$\lim_{k\to\infty} \int_X f_k \, d\mu = \int_X f \, d\mu.$$

2. Suppose that $\{f_k\}_{k=0}^{\infty}$ is almost everywhere nonincreasing, that is, $f_k \geq f_{k+1}$ a.e. Suppose also that there exists an integrable function $g : X \to \overline{\mathbb{R}}$ such that $g \geq f_k$ a.e. for all $k \in \mathbb{N}$. Then, f and f_k are extended integrable for all $k \in \mathbb{N}$, and

$$\lim_{k\to\infty} \int_X f_k \, d\mu = \int_X f \, d\mu.$$

Proof. Since g is integrable, there exists a null set $\widetilde{N} \in \mathfrak{M}$ such that $\{|g| = \infty\} \subset \widetilde{N}$. Now g is $\mathbb{R}$-valued in N^c. We can also select a null set $N \supset \widetilde{N}$ such that the following holds:

- g is measurable on N^c.
- $f_k \to f$ as $k \to \infty$ on N^c.
- For each $k \in \mathbb{N}$, f_k is measurable on N^c, $f_k \leq f_{k+1} \leq f$ on N^c, and $g \leq f_k \leq f$ on N^c.

By Lemma 10.3.22, we know f, $f - g$ are extended integrable on N^c and f_k, $f_k - g$ are extended integrable on N^c for each $k \in \mathbb{N}$. Additionally, we have

$$\int_{N^c} (f - g) \, d\mu = \int_{N^c} f \, d\mu - \int_{N^c} g \, d\mu,$$

and for each $k \in \mathbb{N}$

$$\int_{N^c} (f_k - g) \, d\mu = \int_{N^c} f_k \, d\mu - \int_{N^c} g \, d\mu.$$

Note now $f_k - g$ is measurable function on N^c taking values in $[0, \infty]$. Also, $f_k - g \leq f_{k+1} - g$ on N^c and $f_k - g \to f - g$ pointwise as $k \to \infty$ on N^c. By the basic version of monotone convergence theorem, we have

$$\lim_{k\to\infty} \int_{N^c} (f_k - g) \, d\mu = \int_{N^c} (f - g) \, d\mu.$$

Therefore,

$$\lim_{k\to\infty} \int_{N^c} f_k \, d\mu - \int_{N^c} g \, d\mu = \int_{N^c} f \, d\mu - \int_{N^c} g \, d\mu.$$

However, note that $\int_{N^c} g \, d\mu \in \mathbb{R}$ and it then follows that

$$\lim_{k\to\infty} \int_{N^c} f_k \, d\mu = \int_{N^c} f \, d\mu.$$

Since both f_k and f are extended integrable and N is null, we have

$$\lim_{k\to\infty}\int_X f_k\,d\mu = \int_X f\,d\mu,$$

as desired.

□

Corollary. 1. Let $(X,\mathfrak{M},\mu)$ be a measure space, $f_k : X \to (-\infty,\infty]$ be μ-measurable for all $k\in\mathbb{N}$ and $f_k \geq 0$ a.e. Then,

$$\int_X \sum_{k=0}^{\infty} f_k\,d\mu = \sum_{k=0}^{\infty}\int_X f_k\,d\mu.$$

2. Suppose $(X,\mathfrak{M},\mu)$ is a measure space, $X = \bigcup_{k=0}^{\infty} E_k$ such that $\{E_k\}_{k=0}^{\infty} \subset \mathfrak{M}$ and $\mu(E_k\cap E_j) = 0$ for all $k \neq j$. Given $f : X \to [0,\infty]$ μ-measurable, we then have

$$\int_X f\,d\mu = \sum_{k=0}^{\infty}\int_{E_k} f\,d\mu.$$

Proof. 1. Note that $\operatorname{spt}(f_k^-)$ is in a null set, so each f_k is extended integrable. The same holds for $\sum_{k=0}^{\infty} f_k : X \to [-\infty,\infty]$. On the other hand, the partial sums $\sum_{k=0}^{m} f_k \leq \sum_{k=0}^{m+1} f_k$ a.e. Apply monotone convergence theorem gives the desired equality.

2. Use the first claim on $f_k = f\chi_{E_k}$.

□

Theorem (Fatou's lemma)**.** Let $(X,\mathfrak{M},\mu)$ be a measure space, and suppose that $f_k : X \to \overline{\mathbb{R}}$ are μ-measurable for all $k\in\mathbb{N}$. Suppose that $g : X \to \overline{\mathbb{R}}$ is extended integrable, $\int_X g\,d\mu > -\infty$, and $g \leq f_k$ a.e. for all $k\in\mathbb{N}$. Then the following holds:

1. For each $k\in\mathbb{N}$, f_k is extended integrable.
2. The function $\liminf_{k\to\infty} f_k$ is extended integrable.
3. We have
$$\int_X g\,d\mu \leq \int_X \liminf_{k\to\infty} f_k\,d\mu \leq \liminf_{k\to\infty}\int_X f_k\,d\mu.$$

Proof. Note that $\int_X g\,d\mu > -\infty$ implies g^- is integrable. Write

$$f = \liminf_{k\to\infty} f_k,$$

which is a μ-measurable function. Then, $g \le f_k$ a.e. implies $g \le f$ a.e. as well. It follows that $-f_k \le -g$ and $-f \le -g$. Therefore, $f_k^- \le g^-$ and $f^- \le g^-$. This shows that f_k and f are extended-integrable. Next, note that

$$\int_X g \, d\mu \le \int_X \inf_{j \ge k} f_j \, d\mu \le \int_X f_k \, d\mu.$$

By monotone convergence theorem, we know the middle term converges when k approaches infinity. Taking the liminf, we have

$$\int_X g \, d\mu \le \liminf_{k\to\infty} \int_X \inf_{j\ge k} f_k \, d\mu = \lim_{k\to\infty} \int_X \inf_{j\ge k} f_j \, d\mu = \int_X \liminf_{k\to\infty} f_k \, dx \le \liminf_{k\to\infty} \int_X f_k \, d\mu.$$

□

Theorem (Dominated convergence theorem)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space and suppose $f_k, g_k : X \to \overline{\mathbb{R}}$ μ-measurable for each $k \in \mathbb{N}$. Suppose that $f, g : X \to \overline{\mathbb{R}}$ are such that $f_k \to f$ a.e. and $g_k \to g$ a.e. Suppose further that g_k is integrable and $|f_k| \le g_k$ a.e. for each $k \in \mathbb{N}$. Suppose also g is integrable and that

$$\lim_{k\to\infty} \int_X g_k \, d\mu = \int_X g \, d\mu.$$

Then, f_k is integrable for each $k \in \mathbb{N}$, f is integrable, and

$$\lim_{k\to\infty} \int_X f_k \, d\mu = \int_X f \, d\mu.$$

Moreover, $f_k - f$ is well-defined for all $k \in \mathbb{N}$ outside a null set $N \subset X$, and

$$\lim_{k\to\infty} \int_{N^c} |f_k - f| \, d\mu = 0$$

Proof. We know $|f_k| \le g_k$ a.e. $g_k \to g$ a.e. and $f_k \to f$ a.e. Then, $|f| \le g$ a.e. so f_k and f are integrable. In turn, we can use a previous corollary to pick $N \in \mathfrak{M}$ null such that f_k, f, g_k, g are all $\mathbb{R}$-valued and all assumed inequalities hold on N^c. Then, $|f - f_k| \le g + g_k$ on N^c, and so

$$0 \le g + g_k - |f - f_k| \, .$$

Apply Fatou's lemma, we then have

$$\begin{aligned}
\int_{N^c} 2g \, d\mu &= \int_{N^c} \liminf_{k\to\infty} (g + g_k - |f - f_k|) \, d\mu \\
&\le \liminf_{k\to\infty} \int_{N^c} (g + g_k - |f - f_k|) \, d\mu \\
&= \liminf_{k\to\infty} \int_{N^c} (g + g_k - |f - f_k|) \, d\mu + \liminf_{k\to\infty} \int_{N^c} -(g + g_k) \, d\mu + \int_{N^c} 2g \, d\mu \\
&\le \liminf_{k\to\infty} \int_{N^c} - |f - f_k| \, d\mu + \int_{N^c} 2g \, d\mu.
\end{aligned}$$

It follows that

$$0 \le \limsup_{k\to\infty} \int_{N^c} |f - f_k| \, d\mu = -\liminf_{k\to\infty} \int_{N^c} -|f - f_k| \, d\mu \le 0.$$

Therefore,

$$\lim_{k\to\infty} \int_{N^c} |f - f_k| \, d\mu = 0.$$

Note that f_k and f are integrable, so

$$\left| \int_X f \, d\mu - \int_X f_k \, d\mu \right| = \left| \int_{N^c} f \, d\mu - \int_{N^c} f_k \, d\mu \right| \le \int_{N^c} |f - f_k| \, .$$

This then implies

$$\lim_{k\to\infty} \int_X f_k \, d\mu = \int_X f \, d\mu,$$

and the proof is complete. ☐

Remark. Usually, dominated convergence theorem is applied with $g_k = g$, in which case the assumption $\int_X g_k \, d\mu \to \int_X g \, d\mu$ becomes trivial.

3.4.2 Bochner integration

Lemma. Suppose $(X, \mathfrak{M}, \mu)$ is a measure space and V a normed vector space, and $\phi : X \to V$ simple. Note then $\|\phi\| : X \to [0, \infty)$ is a simple function now. Then, ϕ is a **finite** simple function if and only if $\|\phi\|$ is integrable.

Proof. ($\Longrightarrow$) Suppose ϕ is finite, then $\|\phi\|$ is finite. Then, $\|\phi\|$ is integrable.

($\Longleftarrow$) Suppose $\|\phi\|$ is integrable. We know ϕ is simple, so $\phi(X) \setminus \{0\}$ is a finite set in V. Then, there exists $0 < m \in \mathbb{R}$ such that $\|v\| \ge m$ for all $v \in \phi(X) \setminus \{0\}$. Then,

$$\mu(\operatorname{spt}(\phi)) = \mu(\{x \in X : \|\phi(x)\| > 0\}) = \mu(\{\|\phi\| \ge m\}).$$

By Chebyshev's inequality, we have

$$\mu(\operatorname{spt}(\phi)) \le \frac{1}{m} \int_X \|\phi\| \, d\mu < \infty.$$

This completes the proof. ☐

Lemma. Let $(X, \mathfrak{M}, \mu)$ be a measure space, V be a Banach space, $f : X \to V$ μ-strongly measurable. Suppose that for $j \in \{0, 1\}$, we have $\{\phi_k^j\}_{k=0}^\infty \subset S_{\text{fin}}(X; V)$ such that

$$\lim_{k\to\infty} \int_X \left\| f - \phi_k^j \right\| \, d\mu = 0.$$

Then, $\{\int_X \phi_k^j\}_{k=0}^\infty$ is convergent in V for both $j \in \{0,1\}$ and

$$\lim_{k\to\infty} \int_X \phi_k^0 \, d\mu = \lim_{k\to\infty} \int_X \phi_k^1 \, d\mu.$$

Proof. For $k, m \in \mathbb{N}$, we have

$$\begin{aligned}
\left\| \int_X \phi_m^j \, d\mu - \int_X \phi_k^j \, d\mu \right\| &= \left\| \int_X (\phi_m^j - \phi_k^j) \, d\mu \right\| \\
&\le \int_X \left\| \phi_m^j - \phi_k^j \right\| \, d\mu \\
&\le \int_X \left\| f - \phi_m^j \right\| \, d\mu + \int_X \left\| f - \phi_k^j \right\| \, d\mu.
\end{aligned}$$

This shows that $\{\int_X \phi_k^j\}_{k=0}^\infty$ is Cauchy and hence convergent.

On the other hand,

$$\begin{aligned}
\left\| \int_X \phi_k^0 \, d\mu - \int_X \phi_k^1 \, d\mu \right\| &\le \int_X \left\| \phi_k^0 - \phi_k^1 \right\| \, d\mu \\
&\le \int_X \left\| f - \phi_k^0 \right\| \, d\mu + \int_X \left\| f - \phi_k^1 \right\| \, d\mu \\
&\to 0,
\end{aligned}$$

completing the proof. □

This leads to the following definition for Bochner integration.

Definition. Let $(X, \mathfrak{M}, \mu)$ be a measure space and V a Banach space. A map $f : X \to V$ is (Bochner) integrable if it is strongly μ-measurable and there exists a sequence $\{\phi_n\}_{n=0}^\infty \subset S_{\text{fin}}(X; V)$ such that $\phi_n \to f$ a.e. and

$$\lim_{n\to\infty} \int_X \|f - \phi_n\| \, d\mu = 0,$$

in which case we define

$$\int_X f \, d\mu = \lim_{n\to\infty} \int_X \phi_n \, d\mu \in V.$$

Note that this is well-defined by the previous lemmas.

Theorem (absoulte integrability)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space, V a Banach space, $f : X \to V$. Then, f is integrable if and only if f is μ-measurable and $\|f\| : X \to [0, \infty]$ is integrable. In either case,

$$\left\| \int_X f \, d\mu \right\| \le \int_X \|f\| \, d\mu.$$

Proof. ($\implies$) Suppose f is integrable. This implies that f is strongly μ-measure and in particular μ-measurable. Also, $\|f\| : X \to [0, \infty)$ is μ-measurable. Suppose $\{\phi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; V)$ is such that $\phi_n \to f$ a.e. and $\int_X \|f - \phi_n\| \to 0$. Then,

$$\int_X \|f\| \, d\mu \leq \int_X \|f - \phi_n\| \, d\mu + \int_X \|\phi_n\| \, d\mu < \infty$$

for n sufficiently large. This implies that $\|f\|$ is integrable.

($\impliedby$) Suppose f is μ-measurable and $\int_X \|f\| \, d\mu < \infty$. Then, Pettis theorem gives a sequence $\{\phi_n\}_{n=0}^{\infty} \in S(X; V)$ such that $\phi_n \to f$ a.e. and $\|\phi_n\| \leq 2 \|f\|$. Then,

$$\int_X \|\phi_n\| \, d\mu \leq 2 \int_X \|f\| \, d\mu < \infty.$$

Therefore, $\{\phi_n\}_{n=0}^{\infty}$ is actually a sequence of finite simple functions. This implies that f is actually strongly μ-measurable. On the other hand. $\|f - \phi_n\| \leq 3 \|f\|$, so dominated convergence theorem implies

$$\int_X \|f - \phi_n\| \, d\mu \to 0$$

as $n \to \infty$. By definition, f is now integrable. Moreover,

$$\int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu.$$

It follows then from the dominated convergence theorem that

$$\left\| \int_X f \, d\mu \right\| = \lim_{n \to \infty} \left\| \int_X \phi_n \, d\mu \right\| \leq \lim_{n \to \infty} \int_X \|\phi_n\| \, d\mu = \int_X \|f\| \, d\mu.$$

□

Theorem (dominated convergence theorem for Bochner)**.** Let $(X, \mathfrak{M}, \mu)$ be a measure space, V a Banach space, and suppose $f_n : X \to V$, $g_n : X \to \overline{\mathbb{R}}$ are μ-measurable $n \in \mathbb{N}$. Further suppose $f : X \to V$ and $g : X \to \overline{\mathbb{R}}$ are such that $f_n \to f$ a.e. and $g_n \to g$ a.e. Also, suppose g_n, g are integrable. Finally suppose $\|f_n\| \leq g_n$ a.e. and

$$\lim_{n \to \infty} \int_X g_n \, d\mu = \int_X g \, d\mu.$$

Then, f_n, f are integrable and

$$\lim_{n \to \infty} \int_X \|f_n - f\| \, d\mu = 0,$$

so we also have

$$\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Proof. Since $\|f_n\| \leq g_n$ and $\|f\| \leq g$, we have f_n and f integrable. Note that $\|f - f_n\| \leq g + g_n$ and $g + g_n \to 2g$ as $n \to \infty$. Dominated convergence theorem then implies

$$\lim_{n\to\infty} \int_X \|f - f_n\| \, d\mu = 0,$$

completing the proof. □

Proposition. Let $(X, \mathfrak{M}, \mu)$ be a measure space and V a Banach space over $\mathbb{F}$. Let $f : X \to V$ integrable. The following holds:

1. If W is a Banach space over F and $T \in \mathcal{L}(V, W)$, then $T \circ f : X \to W$ is integrable and

$$\int_X T \circ f \, d\mu = T \int_X f \, d\mu.$$

2. Suppose $g : X \to V$ is integrable, then $\int_X f \, d\mu = \int_X g \, d\mu$ if and only if $\int_X w \circ f \, d\mu = \int_X w \circ g \, d\mu$ for every $w \in V^*$.

Proof. 1. Let $\{\phi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; V)$ such that $\phi_n \to f$ a.e. and $\int_X \|f - \phi_n\| \, d\mu \to 0$. Then we have $T \circ \phi_n \to T \circ f$ a.e. and

$$\int_X \|T \circ f - T \circ \phi_n\| \, d\mu \leq \|T\| \int_X \|f - \phi_n\| \, d\mu \to 0.$$

Therefore, $T \circ f$ is integrable and

$$\int_X T \circ f \, d\mu = \lim_{n\to\infty} \int_X T \circ \phi_n \, d\mu = \lim_{n\to\infty} T \int_X \phi_n \, d\mu = T \int_X f \, d\mu.$$

2. Let $w \in V^*$, then $\int_X f \, d\mu = \int_X g \, d\mu$ clearly implies $\int_X w \circ f \, d\mu = \int_X w \circ g \, d\mu$. On the other hand, if $\int_X w \circ f \, d\mu = \int_X w \circ g \, d\mu$ for all $w \in V^*$, then

$$w \left[\int_X f \, d\mu - \int_X g \, d\mu \right] = 0$$

for all $w \in V^*$. By Hahn-Banach theorem, this implies $\int_X f \, d\mu = \int_X g \, d\mu$.

□

3.5 Products, Fubini-Tonelli, and distribution function

3.5.1 Product measures

Definition (Pre-measure)**.** Let X be a set and $\mathfrak{A}$ be an algebra on X. A map $\gamma : \mathfrak{A} \to [0, \infty]$ is a **pre-measure** if the following is satisfied:

1. $\gamma(\emptyset) = 0$.
2. If $\{A_i\}_{i=0}^\infty \subset \mathfrak{A}$ is disjoint and $\bigcup_{i=0}^\infty A_i \in \mathfrak{A}$, then $\gamma\left(\bigcup_{i=0}^\infty A_i\right) = \sum_{i=0}^\infty \gamma(A_i)$.

Theorem (Pre-measure extension theorem)**.** Let X be a set, $\mathfrak{A}$ is an algebra on X, and γ a pre-measure. Let $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be the outer measure constructed from (X, γ). Denote $\mathfrak{M}$ as the the measurable space and $\mu : \mathfrak{M} \to [0, \infty]$ the corresponding measure. Then the following holds:

1. $\mathfrak{A} \subset \mathfrak{M}$ and $\mu = \gamma$ on $\mathfrak{A}$.
2. Suppose $\mathfrak{N}$ is a σ-algebra on X such that $\mathfrak{A} \subset \mathfrak{N} \subset \mathfrak{M}$, and $\nu : \mathfrak{N} \to [0, \infty]$ is a measure such that $\nu = \gamma$ on $\mathfrak{A}$. Then $\nu \le \mu$ on $\mathfrak{N}$ and $\nu(E) = \mu(E)$ whenever E is σ-finite w.r.t. μ.

   In particular, if X is "γ σ-finite", then $\mu = \nu$ on $\mathfrak{N}$.

Proof. First show $\mu = \gamma$ on $\mathfrak{A}$. It suffices to show that $\mu^* = \gamma$ on $\mathfrak{A}$.

For any $E \in \mathfrak{A}$, we know E is covered by E, so $\mu^* \le \gamma$. On the other hand, let $E \subset \mathfrak{A}$ and $\{A_k\}_{k=0}^\infty \subset \mathfrak{A}$ be a cover of E. Define $B_0 = E \cap A_0 \in \mathfrak{A}$ and $B_k = E \cap (A_k \setminus \bigcup_{i=0}^{k-1} A_k) \in \mathfrak{A}$. Then $\{B_k\}_{k=0}^\infty$ is pairwise disjoint and $\bigcup_{k=0}^\infty B_k = E$. It follows that

$$\gamma(E) = \gamma\left(\bigcup_{k=0}^\infty B_k\right) = \sum_{k=0}^\infty \gamma(B_k) \le \sum_{k=0}^\infty \gamma(A_k).$$

Therefore, $\mu^* = \gamma$ on $\mathfrak{A}$.

Next we show $\mathfrak{A} \subset \mathfrak{M}$. Let $E \in \mathfrak{A}$ be arbitrary and we want to show $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$ for all $A \subset X$. Fix arbitrary $A \subset X$ and $\varepsilon > 0$. Pick $\{A_k\}_{k=0}^\infty \subset \mathfrak{A}$ covering A such that

$$\sum_{k=0}^\infty \gamma(A_k) < \mu^*(A) + \varepsilon.$$

It follows that

$$\begin{aligned}
\mu^*(A \cap E) + \mu^*(A \cap E^c) &\le \mu^*\left(\bigcup_{k=0}^\infty A_k \cap E\right) + \mu^*\left(\bigcup_{k=0}^\infty A_k \cap E^c\right) \\
&\le \sum_{k=0}^\infty \mu^*(A_k \cap E) + \mu^*(A_k \cap E^c) \\
&= \sum_{k=0}^\infty \gamma(A_k \cap E) + \gamma(A_k \cap E^c) \\
&= \sum_{k=0}^\infty \gamma(A_k).
\end{aligned}$$

This implies that E is measurable, completing the proof for the first item.

For the second item, we first show that $\nu \leq \mu$. Let $E \in \mathfrak{N} \subset \mathfrak{M}$ and $\{A_k\}_{k=0}^{\infty} \subset \mathfrak{A}$ that covers E. It follows that

$$\nu(E) \leq \nu\left(\bigcup_{k=0}^{\infty} A_k\right) = \lim_{n\to\infty} \nu\left(\bigcup_{i=0}^{n} A_i\right).$$

Note that $\bigcup_{i=0}^{n} A_i \in \mathfrak{A}$, so $\nu\left(\bigcup_{i=0}^{n} A_i\right) = \mu\left(\bigcup_{i=0}^{n} A_i\right)$. This implies that

$$\nu(E) = \lim_{n\to\infty} \mu\left(\bigcup_{i=0}^{n} A_i\right) = \mu\left(\bigcup_{k=0}^{\infty} A_k\right) \leq \sum_{k=0}^{\infty} \gamma(A_k).$$

Therefore, $\nu \leq \mu$.

Next we show $\nu(E) = \mu(E)$ for $\mu(E) < \infty$. Let $\varepsilon > 0$ and select $\{A_k\}_{k=0}^{\infty} \subset \mathfrak{A}$ covering E such that

$$\sum_{k=0}^{\infty} \gamma(A_k) < \mu^*(E) + \varepsilon = \mu(E) + \varepsilon.$$

Then,

$$\mu\left(\bigcup_{k=0}^{\infty} A_k\right) \leq \sum_{k=0}^{\infty} \gamma(A_k) < \mu(E) + \varepsilon.$$

It follows that $\mu\left(\bigcup_{k=0}^{\infty} A_k \setminus E\right) < \varepsilon$ and thus

$$\mu(E) \leq \mu\left(\bigcup_{k=0}^{\infty} A_k\right) = \nu\left(\bigcup_{k=0}^{\infty} A_k\right) = \nu(E) + \nu\left(\bigcup_{k=0}^{\infty} A_k \setminus E\right) \leq \nu(E) + \varepsilon,$$

where for $\mu\left(\bigcup_{k=0}^{\infty} A_k\right) = \nu\left(\bigcup_{k=0}^{\infty} A_k\right)$ we used the same limit argument as the previous part.

For the case where E is σ-finite, it follows from a similar argument.

$\square$

Theorem (Product measures)**.** Let $2 \leq n \in \mathbb{N}$ and suppose $(X_i, \mathfrak{M}_i, \mu_i)$ is measure space for $1 \leq i \leq n$. Let $X = \prod_i X_i$ and

$$\mathcal{E} = \left\{E = \prod_i E_i : E_i \in \mathfrak{M}_i \text{ for } 1 \leq i \leq n\right\}.$$

The following holds:

1. $\mathfrak{A} = \left\{\bigcup_{k=0}^{K} A^k : \left\{A^k\right\}_k \subset \mathcal{E} \text{ and disjoint}\right\}$ is an algebra.
2. Suppose $\{E^k\}_{k=0}^{\infty} \subset \mathcal{E}$ and $\{F^k\}_{k=0}^{\infty} \subset \mathcal{E}$ are both pairwise disjoint sequences of sets and $\bigcup_{k=0}^{\infty} E^k = \bigcup_{k=0}^{\infty} F^k$, then

$$\sum_{k=0}^{\infty} \prod_{i=1}^{n} \mu_i(E_i^k) = \sum_{k=0}^{\infty} \prod_{i=1}^{n} \mu_i(F_i^k).$$

3. The map $\gamma : \mathfrak{A} \to [0, \infty]$ defined by

$$\gamma\left(\bigcup_{k=0}^{K} \prod_{i=1}^{n} E_i^k\right) = \sum_{k=0}^{K} \prod_{i=1}^{n} \mu_i(E_i^k)$$

is a well-defined pre-measure.

4. If $(X_i, \mathfrak{M}_i, \mu_i)$ is σ-finite, then X is γ σ-finite.

Proof. 1. Since $\emptyset \in \mathfrak{M}_i$ for all $1 \leq i \leq n$, we know $\emptyset \in \mathcal{E}$. Next let $E, F \in \mathcal{E}$ be such that $E = \prod_{i=1}^{n} E_i$ and $F = \prod_{i=1}^{n} F_i$. Then,

$$E \cap F = \prod_{i=1}^{n} (E_i \cap F_i) \in \mathcal{E}.$$

Similarly,

$$E^c = \bigsqcup_{i=1}^{n} \left(E_i^c \times \prod_{j \neq i} E_j\right) \in \mathcal{E}.$$

This shows that $\mathfrak{A}$ is an algebra.

2. Suppose $\bigcup_{k=0}^{\infty} E^k = \bigcup_{k=0}^{\infty} F^k$, then we have

$$\sum_{k=0}^{\infty} \prod_{i=1}^{n} \chi_{E_i^k}(x_i) = \sum_{k=0}^{\infty} \prod_{i=1}^{n} \chi_{F_i^k}(x_i)$$

for all $x = (x_1, \ldots, x_n) \in X$. Now fix $(x_2, \ldots, x_n)$, we then have

$$\sum_{k=0}^{\infty} \chi_{E_1^k}(x_1)\alpha_1^k = \sum_{k=0}^{\infty} \chi_{F_1^k}(x_1)\beta_1^k,$$

where $\alpha_1^k = \prod_{i=2}^{n} \chi_{E_i^k}(x_i)$ and $\beta_1^k = \prod_{i=2}^{n} \chi_{F_i^k}(x_i)$. Using the monotone convergence theorem and integrate both sides, we have

$$\sum_{k=0}^{\infty} \mu_1(E_1)\alpha_1^k = \sum_{k=0}^{\infty} \mu_1(F_1)\beta_1^k.$$

Iterate this argument gives the desired equality.

3. Suppose $\{A_i\}_{i=0}^{\infty} \subset \mathfrak{A}$ disjoint such that $\bigcup_{i=0}^{\infty} A_i \in \mathfrak{A}$. By construction, there exists sequence $\{F^j\}_{j=0}^{J} \subset \mathfrak{A}$ with $J < \infty$ such that $\bigcup_{i=0}^{\infty} A_i = \bigcup_{j=0}^{J} F_j$. Also, $A_i \in \mathfrak{A}$ for each $i \in \mathbb{N}$, so $\bigcup_{i=0}^{\infty} A_i = \bigcup_{k=0}^{\infty} E^k$ where $\{E^k\}_{k=0}^{\infty} \subset \mathcal{E}$ disjoint. It follows that

$$\gamma\left(\bigcup_{i=0}^{\infty} A_i\right) = \gamma\left(\bigcup_{j=0}^{J} F^j\right) = \sum_{j=0}^{J} \prod_{i=1}^{n} \mu_i(F_i^j) = \sum_{k=0}^{\infty} \prod_{i=1}^{n} \mu_i(E_i^k),$$

where the last equality is by item 2. However,

$$\gamma\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{k=0}^{\infty}\prod_{i=1}^{n}\mu_i(E_i^k) = \sum_{i=0}^{\infty}\gamma(A_i).$$

This shows that γ is a pre-measure.

4. For each $1 \leq i \leq n$, there exists $\{S_i\}_{k=0}^{\infty} \subset \mathfrak{M}_i$ such that $S_i^k \subset S_i^{k+1}$, $\bigcup_{k=0}^{\infty} S_i^k = X_i$, and $\mu_i(S_i^k) < \infty$. Consider $\{A^k\}_{k=0}^{\infty}$ where $A^k = \prod_{i=1}^{n} S_i^k$. Note that

$$X = \bigcup_{k=0}^{\infty} A^k \quad \text{and} \quad \gamma(A^k) = \prod_{i=1}^{n}\mu_i\left(S_i^k\right) < \infty.$$

This completes the proof.

□

Corollary. Suppose that $\{(X_i, \mathfrak{M}_i, \mu_i)\}_{i=1}^{n}$ be a sequence of σ-finite measure space. Let $X = \prod_{i=1}^{n} X_i$ be endowed with the product σ-algebra $\bigotimes_{i=1}^{n}\mathfrak{M}_i$. Let $\mathfrak{A}$ and $\gamma : \mathfrak{A} \to [0, \infty]$ be the algebra and pre-measure from the previous theorem. Then, there exists a unique measure $\nu : \bigotimes_{i=1}^{n}\mathfrak{M}_i \to [0, \infty]$ such that $\nu = \gamma$ on $\mathfrak{A}$. Moreover, ν is σ-finite.

Proof. Use the previous theorem and extend the pre-measure. □

3.6 Area formula and change of variable formula

3.6.1 Area formula

We first need to develop a few facts in linear algebra.

Proposition. Let $V_1, \ldots, V_n, W$ be vector space over $\mathbb{F}$ and $T \in L(V_1, \ldots, V_n; W)$. Suppose $x_i^j \in V_i$ for $j = 0, 1$ and $1 \leq i \leq n$. Then,

$$\begin{aligned} T(x_1^0 + x_1^1, \ldots, x_n^0 + x_n^1) &= \sum_{\beta \in B(n)} T(x^{\beta(1)}, \ldots, x^{\beta(n)}) \\ &= \sum_{m=0}^{n}\sum_{\beta \in B_m(n)} T(x_1^{\beta(1)}, \ldots, x_n^{\beta(n)}), \end{aligned}$$

where

$$\begin{aligned} B(n) &= \{\beta : \{1, \ldots, n\} \to \{0, 1\}\}, \\ B_m(n) &= \left\{\beta \in B(n) : \sum \beta(k) = m\right\}. \end{aligned}$$

Proof. Induction on $n \geq 1$. □

Definition. 1. For $1 \leq k \leq n$ we set

$$\mathcal{A}(n,k) = \left\{(\alpha_1, \dots, \alpha_k) \in \{1, \dots, n\}^k : \alpha_1 < \alpha_2 < \dots < a_k\right\}.$$

We also set $\mathcal{A}(n,0) = \{0\}$.

2. For $1 \leq k \leq n$, let $M \in \mathbb{F}^{n\times k}$, $N \in \mathbb{F}^{k\times n}$, $P \in \mathbb{F}^{n\times n}$. For $\alpha \in \mathcal{A}(n,k)$, we set M_α, $N^\alpha, P_\alpha^\alpha \in \mathbb{F}^{k\times k}$ via

$$(M_\alpha)_{i,j} = M_{\alpha_i,j}, \quad (N^\alpha)_{i,j} = N_{i,\alpha_j}, \quad (P_\alpha^\alpha)_{i,j} = P_{\alpha_i,\alpha_j}.$$

Theorem. Let $M \in \mathbb{F}^{n\times n}$ and $Z \in \mathbb{F}$. Then,

$$\det(zI + M) = z^n + \sum_{k=0}^{n-1} z^k \sum_{\alpha\in\mathcal{A}(n,n-k)} \det(M_\alpha^\alpha).$$

Proof. Fix $z \in \mathbb{F}$. Let $x_i^0 = ze_i \in \mathbb{F}^n$ and $x_i^1 = M_i \in \mathbb{F}^n$ be the i-th column of M. Recall that $\det \in L^n(\mathbb{F}^n;\mathbb{F})$. Therefore,

$$\begin{aligned}\det(zI + M) &= \det(x_1^0 + x_1^1, \dots, x_n^0 + x_n^1)\\ &= \sum_{k=0}^{n} \sum_{\beta\in B_k(n)} \det(x_1^{\beta(1)}, \dots, x_n^{\beta(n)})\\ &= z^n + \sum_{k=1}^{n} \sum_{\beta\in B_k(n)} \det(x_1^{\beta(1)}, \dots, x_n^{\beta(n)}).\end{aligned}$$

Now given $1 \leq k \leq n$ and $\beta \in B_k(n)$, we set $\alpha \in \mathcal{A}(n,k)$ to be an increasing enumeration of $\{1 \leq i \leq n : \beta(i) = 1\}$. This gives a bijection from $\mathcal{A}(n,k)$ to $B_k(n)$. On the other hand, if $\beta \in B_k(n)$, then

$$\det(x_1^{\beta(1)}, \dots, x_n^{\beta(n)}) = z^{n-k}\det(M_\alpha^\alpha),$$

for the $\alpha \in \mathcal{A}(n,k)$ that corresponds to the $\beta \in B_k(n)$. This completes the proof. □

Theorem. Let $1 \leq n \leq m$, $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$. The following holds:

1. (Sylvester's formula) $\det(I_m + AB) = \det(I_n + BA)$.
2. (Cauchy-Binet formula) $\det(BA) = \sum_{\alpha\in\mathcal{A}(m,n)} \det A_\alpha \det B^\alpha$.

In particular, if $A^* \in \mathbb{F}^{n\times m}$ given by $A_{ij}^* = \overline{A_{ji}}$, then $\det(A^*A) = \sum_{\alpha\in\mathcal{A}(m,n)} |\det A_\alpha|^2$.

Proof. 1. We have

$$\begin{bmatrix} I_m & A\\ 0_{n\times m} & I_n\end{bmatrix}\begin{bmatrix} I_m & -A\\ B & I_n\end{bmatrix} = \begin{bmatrix} I_m + AB & 0_{m\times n}\\ B & I_n\end{bmatrix}$$

and
$$\begin{bmatrix} I_m & -A \\ B & I_n \end{bmatrix} \begin{bmatrix} I_m & A \\ 0_{n\times m} & I_n \end{bmatrix} = \begin{bmatrix} I_m & 0_{m\times n} \\ B & I_n + BA \end{bmatrix}.$$
It follows that $\det(I_m + AB) = \det(I_n + BA)$.

2. Fix $z \in \mathbb{F} \setminus \{0\}$. Then,
$$\begin{aligned} z^{-m} \det(zI_m + AB) &= \det(I_m + z^{-1}AB) \\ &= \det(I_n + B(z^{-1}A)) \\ &= z^{-n} \det(zI_n + BA). \end{aligned}$$
It follows that $z^n \det(I_m + AB) = z^m \det(I_n + BA)$. By our previous propositions, we have
$$z^{n+m} + \sum_{k=0}^{m-1} z^{k+n} \sum_{\alpha\in\mathcal{A}(m,m-k)} \det(AB)_\alpha^\alpha = z^{n+m} + \sum_{k=0}^{n-1} z^{k+m} \sum_{\alpha\in\mathcal{A}(n,n-k)} \det(BA)_\alpha^\alpha.$$
Consider the coefficients of degree m, we obtain
$$\sum_{\alpha\in A(n,n)} \det(BA)_\alpha^\alpha = \sum_{\alpha\in A(m,n)} \det(AB)_\alpha^\alpha.$$
Note that LHS $= \det BA$ and $(AB)_\alpha^\alpha = A_\alpha B^\alpha$. This completes the proof.

□

Definition (Jacobian map)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be an open set and $f \in C^1(U;\mathbb{R}^m)$ with $n \leq m$. Define the **Jacobian map** $J_f \in C^0(U;[0,\infty))$ by
$$J_f = [\![Df]\!] = \sqrt{\det(Df)^T Df}.$$

Lemma. Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f \in C^1(U;\mathbb{R}^m)$ for some $n \leq m$. Suppose $z \in U$ is such that $Df(z)$ is injective. Then for $0 < \varepsilon < 1$, there exists $B(z,r) \subset U$ such that

1. $f|_{B(z,r)}$ is a Lipschitz injection.
2. If $E \subset B(z,r)$ is Lebesgue measurable, then $f(E) \in \mathfrak{H}^n(\mathbb{R}^m)$ and
$$(1-\varepsilon)^{n+1} \int_E J_f \, d\lambda \leq \mathcal{H}^n(f(E)) \leq (1+\varepsilon)^{n+1} \int_E J_f \, d\lambda.$$

Proof. Define the following $M = Df(z)$, $L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ such that $LM = I_n$, and $g = f \circ L$, so $f = g \circ M$.

Let $0 < \varepsilon < 1$ and pick $r > 0$ such that
$$(1-\varepsilon) \|M(x-y)\| \leq \|f(x) - f(y)\| \leq (1+\varepsilon) \|M(y-x)\| \tag{A}$$

for all $x, y \in B(z, r)$ and

$$(1+\varepsilon)^{-1} J_f(z) \leq J_f(x) \leq (1-\varepsilon)^{-1} J_f(z) \tag{B}$$

for all $x \in B(z, r)$. Note that

$$\mathcal{H}^n(ME) = J_f(z)\lambda(E).$$

Since ML is the identity on range M, equation (A) gives $[g] \leq 1+\varepsilon$ and $[M \circ f^{-1}] \leq (1-\varepsilon)^{-1}$. It follows that

$$\mathcal{H}^n(f(E)) = \mathcal{H}^n(g(ME)) \leq (1+\varepsilon)^n \mathcal{H}^n(ME) = (1+\varepsilon)^n J_f(z)\lambda(E).$$

Also,

$$J_f(z)\lambda(E) = \mathcal{H}^n(ME) = \mathcal{H}^n(M \circ f^{-1}(f(E))) \leq (1-\varepsilon)^{-n} \mathcal{H}^n(f(E)).$$

Now, equation (B) gives

$$J_f(z)\lambda(E) = \int_E J_f(z)\, d\lambda \leq (1+\varepsilon) \int_E J_f \, d\lambda$$

and

$$J_f(z)\lambda(E) = \int_E J_f(z)\, d\lambda \geq (1-\varepsilon) \int_E J_f \, d\lambda.$$

This completes the proof.

□

Definition. Let X be a set equipped with counting measure $\mathcal{H}^0 : \mathcal{P}(X) \to [0, \infty]$. Let Y be a set and $f : X \to Y$. For any $E \subset X$, define $\mathcal{N}_f(\cdot, E) : Y \to [0, \infty]$ by

$$\mathcal{N}_f(y, E) = \mathcal{H}^0\left(E \cap f^{-1}(\{y\})\right) = \mathcal{H}^0\left(\{x \in E : f(x) = y\}\right).$$

Theorem. Let $F \in F_\sigma(\mathbb{R}^n)$ and $f : F \to \mathbb{R}^m$ be locally Lipschitz with $n \leq m$. If $E \subset F$ is Lebesgue measurable, then $\mathcal{N}_f(\cdot, E) : \mathbb{R}^m \to [0, \infty]$ is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable.

HINT: Write $\mathbb{R}^n = \bigcup_{j=0}^\infty Q_{j,k}^\sharp$ as a disjoint union, where $\{Q_{j,k}\}_{j=0}^\infty$ denotes the dyadic cubes of sidelength 2^{-k}, and then study the maps $N_k : \mathbb{R}^m \to [0, \infty]$ given by

$$N_k = \sum_{j=0}^\infty \chi_{f(E \cap Q_{j,k}^\sharp)}.$$

Proof. Following the hint, for each $k \in \mathbb{N}$, write $\mathbb{R}^n = \bigcup_{j=0}^\infty Q_{j,k}^\sharp$ as a disjoint union of dyadic cubes of sidelength 2^{-k}. Claim that for each $y \in \mathbb{R}^m$,

$$\mathcal{N}_f(y, E) = \lim_{k \to \infty} N_k(y) = \lim_{k \to \infty} \chi_{f(E \cap Q_{j,k}^\sharp)}(y).$$

Indeed, if $\mathcal{N}_f(y,E) < \infty$, then there exists $K \in \mathbb{N}$ large enough such that the points are in different cubes for all $k \in K$. On the other hand, if $\mathcal{N}_f(y,E) = \infty$, then for each $n \in \mathbb{N}$, then the same reasoning tells us that there exists K large enough such that $N_K \geq n$. This shows that $\lim_{k\to\infty} N_k(y) = \infty$.

Now since $N_k \to \mathcal{N}_f(\cdot, E)$ pointwise, it suffices to show that N_k is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable for each $k \in \mathbb{N}$. Note that $F \in F_\sigma(\mathbb{R}^n)$, so F is σ-compact. Also, both E and $Q^\sharp_{j,k}$ is Lebesgue measurable, so $E \cap Q^\sharp_{j,k}$ is Lebesgue measurable. This means that $E \cap Q^\sharp_{j,k}$ is also $\mathfrak{H}^n(\mathbb{R}^n)$ measurable. Since f is locally Lipschitz, $f(E \cap Q^\sharp_{j,k})$ is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable. It follows that the characteristic function $\chi_{f(E\cap Q^\sharp_{j,k})}$ is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable for each $j,k \in \mathbb{N}$. Therefore, $N_k = \sum_{j=0}^\infty \chi_{f(E\cap Q^\sharp_{j,k})}$ is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable, as desired. $\square$

Lemma. Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f \in C^1(U;\mathbb{R}^m)$ for $n \leq m$. Suppose $Df(x)$ is injective for all $x \in V$. Then for all $E \subset U$ Lebesgue measurable, and

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E) \, d\mathcal{H}^n.$$

Proof. Let $E \subset U$ be Lebesgue measurable and $0 < \varepsilon < 1$. Using the previous lemma, we can pick $\{B(x_k,r_k)\}_{k=0}^\infty$ such that $B(x_k,r_k) \subset U$, $f : B(x_k,r_k) \to \mathbb{R}^n$ is Lipschitz injection, $E \subset \bigcup_{k=0}^\infty B(x_k,r_k)$, and

$$(1-\varepsilon)^{n+1} \int_F J_f \, d\lambda \leq \mathcal{H}^n(f(F)) \leq (1+\varepsilon)^{n+1} \int_F J_f \, d\lambda$$

for all $F \subset B(x_k,r_k)$.

Let $E_0 = E \cap B(x_0,r_0)$ and for $k > 0$ let $E_k = E \cap B(x_k,r_k) \setminus \bigcup_{j=0}^{k-1} B(x_j,r_j)$. Then $E = \bigsqcup_{k=0}^\infty E_k$. Applying the inequality, we obtain

$$(1-\varepsilon)^{n+1} \int_{E_k} J_f \, d\lambda \leq \mathcal{H}^n(f(E_k)) \leq (1+\varepsilon)^{n+1} \int_{E_k} J_f \, d\lambda.$$

However, since f is injective when restricted to E_k, we have

$$\mathcal{H}^n(f(E_k)) = \int_{f(E_k)} 1 \, d\mathcal{H}^n = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E_k) \, d\mathcal{H}^n.$$

Summing the inequalities, we can then obtain from monotone convergence theorem that

$$(1-\varepsilon)^{n+1} \int_E J_f \, d\lambda \leq \int_{\mathbb{R}^m} \sum_{k=0}^\infty \mathcal{N}_f(\cdot, E_k) \, d\mathcal{H}^n = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E) \, d\mathcal{H}^n \leq (1+\varepsilon)^{n+1} \int_E J_f \, d\lambda.$$

Since this holds for all $\varepsilon > 0$, we have

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E) \, d\mathcal{H}^n.$$

$\square$

Theorem (Sard's theorem, special case)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f \in C^1(U;\mathbb{R}^m)$ for $n \leq m$. Then the set
$$Z = \{x \in U : J_f(x) = 0\}$$
is Lebesgue measurable and $f(Z) \in \mathfrak{H}^n(\mathbb{R}^m)$ and $\mathcal{H}^n(f(Z)) = 0$.

Proof. Note that Z is relatively closed, so it is Lebesgue measurable. It then suffices to show that the outer measure $\mathcal{H}^n(f(Z)) = 0$.

Write $U = \bigcup_{k=0}^\infty Q_k$ where $\{Q_k\}_{k=0}^\infty$ is a sequence of almost disjoint cubes. It suffices to show $\mathcal{H}^n(f(Z_k)) = 0$, where $Z_k = Z \cap Q_k$. Let $0 < \varepsilon < 1$ and let $f_\varepsilon \in C^1(U;\mathbb{R}^{m+n})$ by $f_\varepsilon(x) = (f(x), \varepsilon x)$. Then f_ε is injective, and
$$Df(x) = \begin{bmatrix} Df(x) \\ \varepsilon I_n \end{bmatrix} \in \mathbb{R}^{(m+n)\times n},$$
which is also injective for each $x \in U$. Also,
$$(Df_\varepsilon)^T Df_\varepsilon = \begin{bmatrix} Df^T & \varepsilon I \end{bmatrix} \begin{bmatrix} Df \\ \varepsilon I \end{bmatrix} = (Df)^T Df + \varepsilon^2 I.$$
It follows that
$$\begin{aligned}
J_{f_\varepsilon}^2 &= \det((Df_\varepsilon)^T Df_\varepsilon) \\
&= \det(\varepsilon^2 I + (Df)^T Df) \\
&= \varepsilon^{2n} + \sum_{j=0}^{n-1} \varepsilon^{2j} \sum_{\alpha \in \mathcal{A}(n,n-j)} \det((Df)^T Df)_\alpha^\alpha \\
&= \det(Df)^T Df + \varepsilon^{2n} + \sum_{j=1}^{n-1} \varepsilon^{2j} \sum_{\alpha \in \mathcal{A}(n,n-j)} \det((Df)^T Df)_\alpha^\alpha \\
&\leq J_f^2 + \varepsilon^2 \left[1 + \sum_{j=1}^{n-1} \sum_{\alpha \in \mathcal{A}(n,n-j)} \det((Df)^T Df)_\alpha^\alpha \right].
\end{aligned}$$
Therefore, for $x \in Q_k$, we have $J_{f_\varepsilon}^2(x) \leq J_f^2(x) + \varepsilon^2 C_k$ for a constant $C_k > 0$ depending only on f and $k \in \mathbb{N}$. If $x \in Z_k$, then $x \in Q_k \cap Z$, so $J_{f_\varepsilon}(x) \leq \varepsilon\sqrt{C_k}$. Note that f_ε is injective and $Df_\varepsilon(x)$ are injective for all $x \in Z_k$, the previous lemma gives
$$\mathcal{H}^n(f_\varepsilon(Z_k)) = \int_{Z_k} J_{f_\varepsilon}\, d\lambda \leq \varepsilon\sqrt{C_k}\lambda(Q_k),$$
but $f(Z_k) = \pi_m(f_\varepsilon(Z_k))$ where π_m is the projection map. Therefore,
$$\mathcal{H}^n(f(Z_k)) \leq \mathcal{H}^n(f_\varepsilon(Z_k)) \leq \varepsilon\sqrt{C_k}\lambda(Q_k).$$
This then implies that $\mathcal{H}^n(f(Z_k)) = 0$. $\square$

Theorem (C^1 area formula)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f \in C^1(U;\mathbb{R}^m)$ for $n \leq m$. If $E \subset U$ is Lebesgue measurable, then

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E) \, d\mathcal{H}^n = \int_{f(E)} \mathcal{N}_f(\cdot, E) \, d\mathcal{H}^n.$$

In particular, if f is injective, then

$$\mathcal{H}^n(f(E)) = \int_E J_f \, d\lambda.$$

Proof. Let $Z = \{J_f = 0\}$, which is closed in U. Therefore, $V = U \setminus Z$ is open. Note that $J_f(x) \neq 0$ implies $Df(x)$ injective. Then, previous lemma implies

$$\int_{V \cap E} J_f \, d\lambda = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E \cap V) \, d\mathcal{H}^n.$$

On the other hand,

$$\int_{E \cap Z} J_f \, d\lambda = 0 = \int_{f(E \cap Z)} \mathcal{N}_f(\cdot, E \cap Z) \, d\mathcal{H}^n = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E \cap Z) \, d\mathcal{H}^n.$$

Adding the equality together gives

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E) \, d\mathcal{H}^n.$$

□

3.6.2 Change of variable

Theorem (change of variable, non-injective form)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be open and $f \in C^1(U;\mathbb{R}^m)$ with $n \leq m$. Let $E \subset U$ be measurable. Then the following holds:

1. Suppose $g : E \to [0,\infty]$ is λ-measurable. Then the map

$$\mathbb{R}^m \ni y \mapsto \int_{E \cap f^{-1}(\{y\})} g \, d\mathcal{H}^0 \in [0,\infty] \tag{$*$}$$

   is $\mathcal{H}^n$-measurable, and

$$\int_E g J_f \, d\lambda = \int_{\mathbb{R}^m} \int_{E \cap f^{-1}(\{y\})} g \, d\mathcal{H}^0 \, d\mathcal{H}^n.$$

   In particular, gJ_f is λ-integrable if and only if the map $(*)$ is $\mathcal{H}^n$-integrable.

2. Let $Y \in \{V, \overline{\mathbb{R}}\}$ with V a Banach space. Suppose $g : E \to Y$ is λ-measurable and gJ_f is λ-integrable. Then for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^m$, the restriction $g : E \cap f^{-1}(\{y\}) \to Y$ is $\mathcal{H}^0$-integrable. Moreover, the now Y valued map $(*)$ is $\mathcal{H}^n$-integrable and

$$\int_E g J_f \, d\lambda = \int_{\mathbb{R}^m} \int_{E \cap f^{-1}(\{y\})} g \, d\mathcal{H}^0 \, d\mathcal{H}^n.$$

Example. As an example, say $V \subset \mathbb{R}^n$ and $f : V \to f(V) \in \mathbb{R}^n$ is a diffeomorphism. Then

$$J_f = \sqrt{\det(Df)^T Df} = |\det Df|$$

and

$$\int_{E \cap f^{-1}(\{y\})} g \, d\mathcal{H}^0 = g \circ f^{-1}(y).$$

The theorem then gives

$$\int_E g \, |\det Df| \, d\lambda = \int_{f(E)} g \circ f^{-1} \, d\lambda.$$

This is the usual change of variable formula we encountered before in calculus. △

Proof sketch. 1. We first prove the theorem assuming $g : E \to [0, \infty]$ is Lebesgue measurable. Let $\{\phi_k\}_{k=0}^{\infty}$ be a sequence of simple functions such that $\phi_k \to g$ pointwise as $k \to \infty$. WLOG also assume $\phi_k \leq \phi_{k+1}$. Let

$$\phi_k = \sum_{j=0}^{J_k} \phi_{k,j} \chi_{E_{k,j}}$$

be the canonical representation of ϕ_k.

For $y \in \mathbb{R}^m$, we compute

$$\begin{aligned}\int_{E \cap f^{-1}(\{y\})} \phi_k \, d\mathcal{H}^0 &= \sum_j \phi_{k,j} \mathcal{H}^0(E_{k,j} \cap f^{-1}(\{y\})) \\ &= \sum_j \phi_{k,j} \mathcal{N}_f(y, E_{k,j})\end{aligned}$$

Therefore, the map

$$y \mapsto I_k := \int_{E \cap f^{-1}(\{y\})} \phi_k \, d\mathcal{H}^0$$

is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable. Note that $\phi_k \leq \phi_{k+1}$, so $I_k \leq I_{k+1}$. Monotone convergence theorem then implies that the map

$$y \mapsto I := \int_{E \cap f^{-1}(\{y\})} g \, d\mathcal{H}^0$$

is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable and $I = \lim_{k \to \infty} I_k$

On the other hand,

$$\begin{aligned}\int_E \phi_k J_f \, d\lambda &= \sum_j \phi_{k,j} \int_{E_{k,j}} J_f \, d\lambda \\ &= \sum_j \phi_{k,j} \int_{\mathbb{R}^m} \mathcal{N}_f(\cdot, E_{k,j}) \, d\mathcal{H}^n \\ &= \int_{\mathbb{R}^m} \int_{E \cap f^{-1}(\{y\})} \phi_k \, d\mathcal{H}^0 \, d\mathcal{H}^n(y).\end{aligned}$$

Using monotone convergence theorem again, we obtain

$$\int_E gJ_f\,d\lambda = \int_{\mathbb{R}^m}\int_{E\cap f^{-1}(\{y\})} g\,d\mathcal{H}^0\,d\mathcal{H}^n(y).$$

Therefore, item 1 is proved in the special case. The general case follows by considering null sets and using the more general convergence theorems.

2. To promote from $Y=[0,\infty]$ to $Y=\overline{\mathbb{R}}$ by splitting $g=g^+-g^-$ and applying item 1 to $g^\pm$. Then promote to $Y=\mathbb{C}$ by splitting $g=\operatorname{Re}g+i\operatorname{Im}g$. Finally, promote to V a Banach space over $\mathbb{F}$ as follows: let $w\in V^*$ and consider $w\circ g:E\to\mathbb{F}$. Then show

$$\int w\circ gJ_f\,d\lambda = \iint w\circ g\,d\mathcal{H}^0\,d\mathcal{H}^n$$

for all $w\in V^*$. This will then give the desired result.

$\square$

Theorem (Change of variable, injective form)**.** Let $\emptyset\neq U\subset\mathbb{R}^n$ be open and $f\in C^1(U;\mathbb{R}^m)$ with $n\leq m$. Let $Z=\{J_f=0\}$ and suppose that $\lambda(Z)=0$. Suppose that $f:U\setminus Z\to f(U\setminus Z)$ is injective. Let $E\subset U$ be Lebesgue measurable and $g:f(E)\to Y$, where $Y\in\{V,\overline{\mathbb{R}}\}$ with V a Banach space. Then the following holds:

1. $f(E)\in\mathfrak{H}^n(\mathbb{R}^m)$.
2. g is $\mathcal{H}^n$-measurable if and only if $g\circ f$ is λ-measurable.
3. g is $\mathcal{H}^n$-integrable on $f(E)$ if and only if $g\circ fJ_f$ is λ-integrable on E. In either case,

$$\int_E g\circ fJ_f\,d\lambda = \int_{f(E)} g\,d\mathcal{H}^n.$$

Theorem (change of variable, local injective form)**.** Let $\emptyset\neq U\subset\mathbb{R}^n$ be open, $f\in C^1(U;\mathbb{R}^m)$ for $n\leq m$. Suppose $E\subset U$ is Lebesgue measurable such that $E^\circ\neq\emptyset$ and $\lambda(\partial E\cap U)=\lambda(Z\cap E)=0$ where $Z=\{J_f=0\}$. Further suppose the restriction $f:E^\circ\to f(E^\circ)$ is injective. Finally let $g:f(E)\to Y$, where $Y\in\{V,\overline{\mathbb{R}}\}$ with V a Banach space. Then the following holds:

1. $f(E)\in\mathfrak{H}^n(\mathbb{R}^m)$.
2. g is $\mathcal{H}^n$-measurable if and only if $g\circ f$ is λ-measurable.
3. g is $\mathcal{H}^n$-integrable on $f(E)$ if and only if $g\circ fJ_f$ is λ-integrable on E. In either case,

$$\int_E g\circ fJ_f\,d\lambda = \int_{f(E)} g\,d\mathcal{H}^n.$$

Proof sketch. Apply the previous theorem to see that

$$\int_{E^\circ} g \circ f J_f \, d\lambda = \int_{f(E^\circ)} g \, d\mathcal{H}^n.$$

However, f being C^1 implies that it is locally Lipschitz, so $\mathcal{H}^n(\partial E \cap E) = 0$. Therefore,

$$\int_{\partial E \cap E} g \circ f J_f \, d\lambda = 0 = \int_{f(\partial E \cap E)} g \, d\mathcal{H}^n.$$

□

Example. Let $E = [0, \infty) \times [0, 2\pi] \subset \mathbb{R}^2$. Let $\rho : \mathbb{R}^2 \to \mathbb{R}^2$ via

$$\rho(r, \theta) = (r \cos\theta, r \sin\theta).$$

Then ρ is a smooth function. Note

$$E^\circ = (0, \infty) \times (0, 2\pi)$$

and the restriction $\rho : E^\circ \to \rho(E^\circ)$ is injective. Additionally,

$$\partial E = (\{0\} \times [0, 2\pi]) \cup ([0, \infty) \times \{0\}) \cup ([0, \infty) \times \{2\pi\})$$

is null. We also know that $|\det D\rho(r, \theta)| = r$, so $Z = \{J_\rho = 0\}$ is null.

Note that $\rho(E^\circ) = \mathbb{R}^2 \setminus N$ where N is null. Then, $f : \mathbb{R}^2 \to V$ is integrable if and only if $f \circ \rho : E^\circ \to V$ is integrable, and in either case,

$$\begin{aligned}
\int_{\mathbb{R}^2} f \, d\lambda &= \int_{E^\circ} f \circ \rho J_\rho \, d\lambda \\
&= \int_0^\infty \int_0^{2\pi} f(r \cos\theta, r \sin\theta) r \, d\theta \, dr \\
&= \int_0^{2\pi} \int_0^\infty f(r \cos\theta, r \sin\theta) r \, dr \, d\theta.
\end{aligned}$$

△

Example. Let $E = [0, \infty) \times [0, 2\pi] \times [0, \pi] \subset \mathbb{R}^3$ and $\zeta : \mathbb{R}^3 \to \mathbb{R}^3$ via

$$\zeta(r, \theta, \phi) = (r \cos\theta \sin\phi, r \sin\theta \sin\phi, r \cos\phi).$$

This is 3D spherical coordinates. Similar to the previous example, $\zeta : E^\circ \to \zeta(E^\circ)$ is injective and $\mathbb{R}^3 \setminus \zeta(E^\circ)$ is null. In addition,

$$J_\zeta(r, \theta, \phi) = |\det D\zeta(r, \theta, \phi)| = r^2 \sin\phi.$$

Then for $f : \mathbb{R}^3 \to V$ we have f is integrable if and only if $f \circ \zeta : E^\circ \to V$ is integrable, and in either case,

$$\begin{aligned}\int_{\mathbb{R}^3} f \, d\lambda &= \int_{E^\circ} f \circ \zeta J_\zeta \, d\lambda \\ &= \int_0^\infty \int_0^{2\pi} \int_0^\pi f \circ \zeta(r, \theta, \phi) r^2 \sin\phi \, d\phi \, d\theta \, dr.\end{aligned}$$

△

Example (Hyperspherical coordinates)**.** Let $n \geq 3$ and $E = [0, \infty) \times [0, 2\pi] \times [0, \pi]^{n-2} \in \mathbb{R}^n$. Let $\Psi : E \to \mathbb{R}^n$ via

$$\begin{aligned}\Psi_n &= r\cos(\phi_{n-2}) \\ \Psi_{n-1} &= r\sin(\phi_{n-2})\cos(\phi_{n-3}) \\ &\cdots \\ \Psi_3 &= r\sin(\phi_{n-2})\sin(\phi_{n-3})\cdots\sin(\phi_2)\cos(\phi_1) \\ \Psi_2 &= r\sin(\phi_{n-2})\sin(\phi_{n-3})\cdots\sin(\phi_2)\sin(\phi_1)\sin(\theta) \\ \Psi_1 &= r\sin(\phi_{n-2})\sin(\phi_{n-3})\cdots\sin(\phi_2)\sin(\phi_1)\cos(\theta)\end{aligned}$$

An induction argument shows that

$$\det D\Psi = r^{n-1}\sin(\phi_1)\sin^2(\phi_2)\ldots\sin^{n-2}(\phi_{n-2}) \geq 0.$$

Then for $f : \mathbb{R}^n \to V$ we have

$$\int_{\mathbb{R}^n} f \, d\lambda = \int_E f \circ \Psi J_\Psi$$

△

3.7 The Lebesgue-Bochner L^p spaces

In this subsection, $(X, \mathfrak{M}, \mu)$ will denote a measure space and V will denote a Banach space.

Definition.

1. For μ-measurable f, for $1 \leq p < \infty$ define $\|p\|_\infty$ via

$$\|f\|_p = \left[\int_X \|f\|^p \, d\mu\right]^{\frac{1}{p}},$$

and for $p = \infty$ define $\|f\|_\infty$ via

$$\|f\|_\infty = \inf\{s \in [0, \infty] : f \leq s \text{ a.e.}\}.$$

2. Define

$$\mathrm{Leb}^p_\mu(X; V) = \{f : X \to V,\ f \text{ is } \mu\text{-measurable and } \|f\|_p < \infty\}$$

to be the set of p-integrable functions.

Example. 1. If $X \in \{\mathbb{N}, \mathbb{Z}\}$ with μ being the counting measure, then $\ell^p(X;V) = \mathrm{Leb}^p_\mu(X;V)$.

2. For all $p \in [1,\infty]$, we have $S_{\text{fin}}(X;V) \subset \mathrm{Leb}^p_\mu(X;V)$.

△

Theorem (Hölder's inequalities)**.** Suppose $p, q, r \in [1,\infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, then the following holds:

1. If $f \in \mathrm{Leb}^p_\mu(X;\mathbb{F})$, $g \in \mathrm{Leb}^q_\mu(X;\mathbb{F})$, then $fg \in \mathrm{Leb}^r_\mu(X;\mathbb{F})$ and
$$\|fg\|_r \leq \|f\|_p \|g\|_q .$$
In particular, if $r = 1$, then
$$\int_X |fg| \, d\mu \leq \|f\|_p \|g\|_q .$$

2. If V_1, V_2, W are Banach spaces and $T \in \mathcal{L}(V_1, V_2; W)$ then
$$\|T(f,g)\| \leq \|T\|_{\mathcal{L}} \|f\|_p \|g\|_q .$$

Theorem (Minkowski's inequalities)**.** Let $p \in [1,\infty]$ and $f, g \in \mathrm{Leb}^p_\mu(X;V)$, then
$$\|f+g\|_p \leq \|f\|_p + \|g\|_p .$$

Corollary. $\mathrm{Leb}^p_\mu(X;V)$ is a vector space for $p \in [1,\infty]$, and $\|\cdot\|_p$ is a semi-norm. Also, $\|f\|_p = 0$ if and only if $f = 0$ a.e.

Theorem. Let $f \in \mathrm{Leb}^p_\mu(X;V)$. For any $\varepsilon > 0$, there exists $\phi \in S_{\text{fin}}(X;V) \subset \mathrm{Leb}^p_\mu(X;V)$ such that $\|\phi - f\|_p < \varepsilon$.

Proof sketch. By Pettis theorem, we can select $\{\psi_n\}_{n=0}^\infty \subset S(X;V)$ such that $\|\psi_n\| \leq 2\|f\|$ on X and $\psi_n \to f$ a.e. We then have $\|\psi_n\|^p \leq 2^p \|f\|$, so $\psi_n \in S_{\text{fin}}(X;V)$. Note that $\|\psi_n - f\| \leq 3\|f\|$ on X. Dominated convergence theorem then implies
$$\lim_{n\to\infty} \int_X \|\psi_n - f\| \, d\mu = \int_X \lim_{n\to\infty} \|\psi_n - f\| \, d\mu = 0.$$
□

Next we define L^p space by taking the qoutient of $\mathrm{Leb}^p_\mu(X;V)$ under equivalence relation $\sim$, where $f \sim g$ if $f = g$ a.e. Once we show $L^p(X;V)$ is complete and $S_{\text{fin}}(X;V)$ is denes in $L^p(X;V)$, we will know the completion of $S_{\text{fin}}(X;V)$ is $L^p(X;V)$.

Definition (L^p spaces)**.** For $f, g \in \mathrm{Leb}^p_\mu$, let $f \sim g$ if $f = g$ a.e. Then $\sim$ is a equivalence relation. Let $[f]$ be the equivalence class of f. Define
$$L^p_\mu(X;V) = \{[f] : f \in \mathrm{Leb}^p_\mu(X;V)\}$$
with $\alpha[f] = [\alpha f]$, $[f] + [g] = [f+g]$, and $\|[f]\|_{L^p} = \|f\|_p$. Then it is easy to see that all the operations are well-defined and L^p_μ is a vector space.

Theorem (Riesz-Fisher)**.** $L^p_\mu(X;V)$ is a Banach space for $p \in [1,\infty]$.

Proof. Let $\{f_k\}_{k=0}^\infty \in L^p$ such that $\sum_{k=0}^\infty \|f_k\|_p < \infty$. We want to show $\sum_{k=0}^\infty f_k$ converges in L^p.

First suppose $p < \infty$. Let

$$g_m = \sum_{k=0}^m \|f_k\|_V \in L^p(X;\mathbb{F}).$$

Then, there exists $S \in \mathbb{R}$ such that

$$\|g_m\|_{L^p} \leq \sum_{k=0}^m \|\|f_k\|_V\|_p = \sum_{k=0}^m \|f_k\|_p < S.$$

for all $m \in \mathbb{N}$. It follows that

$$\int_X |g_m|^p \; d\mu \leq S^p.$$

Let $g = \lim_{m\to\infty} g_m$. Then $g(x) \neq \infty$ a.e. and by monotone convergence theorem, we have

$$\int_X |g|^p \; d\mu = \lim_{n\to\infty} |g_m|^p \; d\mu \leq S^p.$$

Define $F : X \to V$ in the following way:

$$F(x) = \begin{cases} 0 & \text{if } g(x) = \infty, \\ \sum_{k=0}^\infty f_k(x) & \text{if } g(x) \neq \infty. \end{cases}$$

Then $\|F - \sum_{k=0}^m f_k\|_V^p \leq g^p$ a.e. and $\lim_{m\to\infty} \|F - \sum_{k=0}^m f_k\| = 0$ a.e. It then follows that

$$\lim_{m\to\infty} \left\| F - \sum_{k=0}^m f_k \right\|_{L^p} = \lim_{m\to\infty} \int_X \left\| F - \sum_{k=0}^m f_k \right\|_V^p \; d\mu = 0,$$

where the last equality is by dominated convergence theorem. □

4 Manifolds in $\mathbb{R}^n$, differential forms, Stokes-Cartan

We know from the area formula that we can integrate on certain m-dimensional subsets of $\mathbb{R}^n$ ($m \leq n$). We also know through the fundamental theorem of calculus that there exists a beautiful and useful connection between integral and differential calculus in 1D. The goal is then to develop some form of differential calculus on sets we can integrate over in $\mathbb{R}^n$ and to generalize FTC.

There are a few caveats:

1. We will only develop the "extrinsic theory" of manifolds in $\mathbb{R}^n$. This is opposed to the modern "intrinsic" perspective, which does not rely on containment in $\mathbb{R}^n$. However, there are deep theorems in modern manifold theory (Nash[1], Whitney) that show "intrinsic if and only if extrinsic"
2. A lot of the theory we develop is front-loading for the Stokes-Cartan theorem.
3. We are **NOT** doing differential geomtry.
4. All of this works with Banach space replacing $\mathbb{R}^n$. We are working in $\mathbb{R}^n$ to make things easier and avoid some slight subtleties.

4.1 Manifolds in $\mathbb{R}^n$

4.1.1 Definitions and basics

Definition. Let $m, n \in \mathbb{N}$ and $1 \leq m \leq n$ and let $1 \leq k \leq \infty$. Let $M \subset \mathbb{R}^n$.

1. We say M admits C^k local m-coordinates at $z \in M$ if the following holds:
   (a) There exists a set $U \subset \mathbb{R}^n$ open with $z \in U$.
   (b) There exists $\emptyset \neq V \subset \mathbb{R}^m$ open and **homeomorphism** $\phi : V \to \phi(V) = M \cap U$.
   (c) $\phi \in C^k(V; \mathbb{R}^n)$ and $D\phi(x) \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) \simeq \mathbb{R}^{n \times m}$ has rank m for all $x \in V$.

   The triple (U, V, ϕ) is called a C^k m-coordinate **chart**.
2. We say M is an m-dimensional C^k **manifold** if $M \neq \emptyset$ and M admits local C^k m-coordinates at each $z \in M$. If M is an m-dimensional manifold, we say M is an m-manifold.
3. An **atlas** on M is a collection $\mathcal{A} = \{(U_\alpha, V_\alpha, \phi_\alpha)\}_{\alpha \in A}$ for $A \neq \emptyset$ some index set such that $(U_\alpha, V_\alpha, \phi_\alpha)$ is C^k local m-coordinates for all $\alpha \in A$ and $M \subset \bigcup_\alpha U_\alpha$.

Remark. 1. From time to time it is useful to replace $\mathbb{R}^n$ and $\mathbb{R}^m$ with generic real normed vector space of dimension n and m respectively. This changes nothing in the definition.

2. The definition fails to define 0-dimensional manifolds. We make the convention that a 0-manifold is a nonempty discrete subset of $\mathbb{R}^n$, where by discrete we mean every subset of it is both open and closed. We also consider all 0-manifold to be C^∞.

[1] undergraduate student in Carnegie Mellon University

Example. 1. Let $\emptyset \subset U \subset \mathbb{R}^n$ open, then U is a C^∞ n-manifold.

2. Let $w + V = \{w + V : v \in V\}$ for $w \in \mathbb{R}^n$ and $V \subset \mathbb{R}^n$ a subset of dimension m. Indeed, let $\{v_1, \ldots, v_m\} \subset V$ be a basis and consider $\mathbb{R}^m \ni x \mapsto w + \sum_j x_j v_j \in w + V$.

3. Suppose $M \subset \mathbb{R}^n$ is a C^k m-manifold. Suppose $U \subset \mathbb{R}^n$ is open and $M \cap U \neq \emptyset$. Then $M \cap U$ is a C^k m-manifold.

4. Suppose $M_1, M_2 \subset \mathbb{R}^n$ are C^k m-manifolds. Suppose $|x - y| \geq \varepsilon > 0$ for all $x \in M_1$ and $y \in M_2$. Then $M_1 \cup M_2$ is a C^k m-manifold.

5. Suppose $M_1 \subset \mathbb{R}^{n_1}$ and $M_2 \subset \mathbb{R}^{n_2}$ are C^k m_1-manifold and C^k m_2-manifold respectively. Then $M_1 \times M_2 \subset \mathbb{R}^{n_1+n_2}$ is a C^k $(m_1 + m_2)$-manifold.

6. Let $\phi : \mathbb{R} \times (-\frac{1}{2}, \frac{1}{2}) \to \mathbb{R}^3$ via

$$\phi(\theta, t) = (\cos\theta, \sin\theta, t).$$

Then the image M is a cylinder in $\mathbb{R}^3$. Note that ϕ is smooth and $D\phi$ has rank 2 everywhere. We can build an atlas on M via $\mathcal{A} = \{(U_1, V_2, \phi), (U_2, V_2, \phi)\}$ for $V_1 = (0, 2\pi) \times (-\frac{1}{2}, \frac{1}{2})$ and $V_2 = (\pi, 3\pi) \times (-\frac{1}{2}, \frac{1}{2})$. After we figure out the corresponding U_1 and U_2, we can see that M is a C^∞ 2-manifold.

$\triangle$

Theorem (level sets as manifolds)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f \in C^k(U; \mathbb{R}^r)$ for $1 \leq r < n$. Suppose $w_0 \in f(U)$ and

$$M = \{z \in U : f(z) = w_0 \text{ and } Df(z) \text{ had rank } r\} \neq \emptyset.$$

Then M is a C^k $(n - r)$-manifold.

Proof. Let $z \in M$. Then, $Df(z)$ is surjective, so we can apply the **right inverse function theorem**. This provides open sets $W \subset \mathbb{R}^r$ and $V \subset \mathbb{R}^{n-r}$, and a C^k diffeomorphism $F : W \times V \to F(W \times V) \subset U$, such that $w_0 \in W$ and $f(F(w, v)) = w$ for all $(w, v) \in W \times V$. Let $\phi : V \to U$ via $\phi(v) = F(w_0, v)$. Then

$$f(\phi(v)) = f(F(w_0, v)) = w_0.$$

Meanwhile, the right inverse function theorem guarantees that $D\phi$ has rank $n - r$ everywhere. Therefore, $\phi : V \to M \cap U$ is the desired homeomorphism to show that M is a C^k $(n - r)$-manifold. $\square$

Example. 1. Let $S \in \mathbb{R}^{n \times n}$ be symmetric with at least one positive eigenvalue. Consider

$$M = \{x \in \mathbb{R}^n : Sx \cdot x = 1\} \neq \emptyset.$$

Consider $f : \mathbb{R}^n \to \mathbb{R}$ via $f(x) = Sx \cdot x$, which is C^∞. Note that

$$Df(x)h = \nabla f(x) \cdot h = 2Sx \cdot h,$$

and for $x \in M$, we have $Df(x)x = 2 \neq 0$. Therefore, $Df(x)$ has rank 1 for all $x \in M$. The previous theorem then tells us M is a C^∞ $(n-1)$-manifold in $\mathbb{R}^n$.

(a) If $S = I$, then $M = \{x : |x|^2 = 1\} = S^{n-1}$ is a sphere and it is a C^∞ $(n-1)$-manifold in $\mathbb{R}^n$.

(b) If $S = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_i > 0$, then M is an ellipsoid and it is a C^∞ $(n-1)$-manifold in $\mathbb{R}^n$.

(c) If $S = \operatorname{diag}(-\lambda_1, \lambda_2, \dots, \lambda_n)$ with $\lambda_i > 0$, then M is a hyperboloid and it is a C^∞ $(n-1)$-manifold in $\mathbb{R}^n$.

2. Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f \in C^k(U; \mathbb{R}^m)$, and

$$\Gamma(f) = \{(X, f(x)) \in \mathbb{R}^{n+m} : x \in U\}.$$

Let $F : U \times \mathbb{R}^m \to \mathbb{R}^m$ via $F(x, y) = y - f(x)$. Then $(x, y) \in \Gamma(f)$ if and only if $F(x, y) = 0$. However, F is C^k and $DF(x, y)(h, k) = k - Df(x)h$. The k term implies that the derivative has rank m. The previous theorem then implies that $\Gamma(f)$ is a C^k n-manifold in $\mathbb{R}^{n+m}$.

3. Note that $\mathrm{GL}(n) \subset \mathbb{R}^{n\times n}$ is open and hence a C^∞ n^2-manifold.

4. Let $\mathrm{SL}(n) = \{M \in \mathrm{GL}(n) : \det M = 1\}$, then $\det : \mathrm{GL}(n) \to \mathbb{R}$ is C^∞ and

$$D\det(M)(N) = (\det M)\operatorname{tr}(M^{-1}N).$$

Therefore, for $M \in \mathrm{SL}(n)$, $D\det(M)M = \operatorname{tr}(M^{-1}M) = n \neq 0$. This implies that $D\det(M)$ has rank 1 at each $M \in \mathrm{SL}(n)$. This then shows that $\mathrm{SL}(n)$ is a C^∞ (n^2-1)-manifold in $\mathbb{R}^{n\times n}$.

5. The set of orthogonal matrices $O(n)$ is a C^∞ $\frac{n(n-1)}{2}$-manifold in $\mathbb{R}^{n\times n}$. Note that

$$O(n) = \{M \in \mathbb{R}^{n\times n} : M^T M = I\}.$$

Consider the map $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ via $f(M) = M^T M - I$, which is smooth and

$$Df(M)(N) = M^T N + N^T M = M^T N + (M^T N)^T.$$

Therefore, $\ker Df(M) = M^{-T}\mathbb{R}^{n\times n}_{\mathrm{asym}} = M\mathbb{R}^{n\times n}_{\mathrm{asym}}$ and hence $\dim\ker Df(M) = \dim \mathbb{R}^{n\times n}_{\mathrm{asym}} = \frac{n(n-1)}{2}$. It follows that

$$\operatorname{rank} Df(M) = \frac{n(n+1)}{2} = \dim \mathbb{R}^{n\times n}_{\mathrm{sym}}.$$

Therefore, $O(n)$ is a C^∞ $\frac{n(n-1)}{2}$-manifold.

As a side note, $\mathbb{R}^{n\times n} = \mathbb{R}^{n\times n}_{\mathrm{sym}} \oplus \mathbb{R}^{n\times n}_{\mathrm{asym}}$ with the column product $M : N = M_{ij}N_{ij}$ as the inner product.

△

Theorem. Let $m, n \in \mathbb{N}$ with $1 \leq m \leq n$, $1 \leq k \leq \infty$. Let $M \in \mathbb{R}^n$, $z \in M$, then the following are equivalent:

1. M admits C^k local m-coordinates at z, say (U, V, ϕ).
2. There exists an open $W \in \mathbb{R}^n$ and C^k diffeomorphism $F : W \to F(W) \subset \mathbb{R}^n$ such that

$$F(M \cap W) = F(W) \cap \{x \in \mathbb{R}^n : x_l = 0 \text{ for } m+1 \leq k \leq n\}.$$

In either case, ϕ and F are related via $\phi^{-1} = (F_1, \dots, F_n)$.

This is in fact the **flattening map** we encounted in the differential calculus section, and the proof for the theorem follows immediately.

Aside from flattening maps, we have another formulation for manifolds and local coordinates.

Theorem. Let $1 \leq m \leq n-1$ and $1 \leq k \leq \infty$. Suppose that $M \subset \mathbb{R}^n$ and $z \in M$. Then the following are equivalent.

1. M admits local C^k m-coordinates at z.
2. M is the (permuted) graph of a C^k function near z, i.e. there exist
   (a) a permutation matrix $P \in \mathbb{R}^{n \times n}$,
   (b) open sets $\emptyset \neq V \subset \mathbb{R}^m$ and $U \subset \mathbb{R}^n$ with $z \in U$,
   (c) a C^k function $f : V \to \mathbb{R}^{n-m}$,

   such that $U \cap M = P\Gamma(f)$, where recall that

$$\Gamma(f) = \{(x, f(x)) : x \in V\} \subset V \times \mathbb{R}^{n-m} \subset \mathbb{R}^n$$

   is the graph of f.

Note that this means we have three distinct but closely related ways of thinking about manifolds locally: (1) as being parameterized by the coordinate maps; (2) as the level sets of the flattening maps; (3) as the permuted graph of some function.

Proof. (1) $\Longrightarrow$ (2). Suppose M admits local C^k coordinates (U, V, ϕ) at z. Then there exists open set $W \in \mathbb{R}^n$ with $z \in W$ and a flattening map $F : W \to \mathbb{R}^n$ such that F is a C^k diffeomorphism onto its image and

$$F(W \cap M) = F(W) \cap \{x \in \mathbb{R}^n : x_l = 0 \text{ for } m+1 \leq l \leq n\}.$$

It follows that

$$W \cap M = \{x \in W : F_l(x) = 0 \text{ for } m+1 \leq l \leq n\} \subset \mathbb{R}^n.$$

Let $\widetilde{F} : \mathbb{R}^n \to \mathbb{R}^{n-m}$ via $\widetilde{F} = (F_{m+1}, \ldots, F_n)$. Since F is a diffeomorphism, $DF(z)$ is an isomophism. In particular,

$$D\widetilde{F} = \begin{bmatrix} \partial_1 F_{m+1} & \partial_2 F_{m+1} & \cdots & \partial_n F_{m+1} \\ \partial_1 F_{m+2} & \partial_2 F_{m+2} & \cdots & \partial_n F_{m+2} \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 F_n & \partial_2 F_n & \cdots & \partial_n F_n \end{bmatrix}$$

evaluated at z contains a invertible $(n-m) \times (n-m)$ submatrix. First assume now the bottom-right $(n-m) \times (n-m)$ submatrix of $D\widetilde{F}(z)$ is invertible. Note that $\mathbb{R}^m \times \mathbb{R}^{n-m} \simeq \mathbb{R}^n$ so $\widetilde{F}$ can be treated as a function from $\mathbb{R}^m \times \mathbb{R}^{n-m}$ to $\mathbb{R}^n$. By our assumption, $D_2\widetilde{F}$ is invertible, where D_2 is taking the partial derivative with respect to the $\mathbb{R}^{n-m}$ dimension. Let $\pi_1 : \mathbb{R}^n \to \mathbb{R}^m$ and $\pi_2 : \mathbb{R}^n \to \mathbb{R}^{n-m}$ denote the projection onto the first m components and the last $n-m$ components respectively. Then the implicit function theorem applies, and gives open sets $V \subset \mathbb{R}^m$, $S \in \mathbb{R}^{m-n}$ with $\pi_1 z \in V$ and $\pi_2 z \in S$ and a C^k function $f : V \to \mathbb{R}^{n-m}$ such that $U = V \times S \subset W$, and $\widetilde{F}(x, y) = 0$ if and only if $y = f(x)$ for all $(x, y) \in V \times S$. It follows that U is open, and $U \cap M = \Gamma(f)$, as desired.

In the general case, $D\widetilde{F}$ still need to contain a invertible $(n-m) \times (n-m)$ submatrix. Apply the same argument but with the $\mathbb{R}^{n-m}$ dimension align with the invertible submatrix. Then, we know there exists a permutation matrix $P \in \mathbb{R}^{n \times n}$ such that $U \cap M = P\Gamma(f)$, as desired.

(2) $\implies$ (1). Suppose the assumption in item 2 holds and let $z \in M$. By assumption there exists an open set $U \subset \mathbb{R}^n$ with $z \in U$ and $\emptyset \neq V \subset \mathbb{R}^m$, and a C^k function $f : V \to \mathbb{R}^{n-m}$ and a permutation matrix $P \in \mathbb{R}^{n \times n}$ such that $U \cap M = P\Gamma(f)$. Consider now the function $\phi : V \to \mathbb{R}^n$ by $\phi(x) = P(x, f(x))$. It is clear that ϕ is C^k and $\phi(V) = P\Gamma(f) = U \cap M$. Moreover, we have

$$D\phi(x)(h) = P(h, Df(x)h).$$

Note that P is a permutation matrix so the h term guarantees that $D\phi(x)$ is at least rank m. However, $V \subset \mathbb{R}^m$ so $D\phi(x)$ is rank m for all $x \in V$. This shows that (U, V, ϕ) is a C^k m-coordinate chart at z. Since z is arbitrary, M admits a local C^k m-coordinate at z. □

Why are these formulations important? There are two reasons.

1. This unifies the perspective on m-manifolds. We saw already that

$$M = \{x : f(x) = y \text{ and } Df(x) \text{ had full rank}\}.$$

   is a manifold. The flattening map formulation tells us that this is the only example we need locally. That is, all manifolds locally are locally a level set of some functions.

2. They are useful for certain technical points, for example the theorem below that tells us about the connection between different local coordinate charts around the same point in the manifold.

Theorem. Suppose (U_0, V_0, ϕ_0) and (U_1, V_1, ϕ_1) are C^k local m-coordinates at $z \in M$. Suppose also $U_0 \cap U_1 \neq \emptyset$ and $z \in U_0 \cap U_1 \cap M$. Then the transition map $\zeta : \phi_0^{-1}(U_0 \cap U_1 \cap M) \to \phi_1^{-1}(U_0 \cap U_1 \cap M)$ is a C^k diffeomorphism.

Proof. Both sets are open because $\phi_i : V_i \to \phi_i(V_i) = M \cap U_i$ is homeomorphism for each $i \in \{0, 1\}$. Let $\widetilde{V_i} = \phi_i^{-1}(U_0 \cap U_1 \cap M)$ for each $i \in \{0, 1\}$ and fix $x \in \widetilde{V_0}$. By the previous theorem, we can find a flattening map $F : W \to F(W)$ such that F is a C^k diffeomorphism and $\phi_1^{-1} = (F_1, \ldots, F_m) = \widetilde{F}$. Therefore, $\zeta = \phi_1^{-1} \circ \phi_0 = \widetilde{F} \circ \phi_0$ near x. However $x \in \widetilde{V_0}$ is arbitrary, so ζ is a C^k diffeomorphism. □

4.1.2 Hausdorff measure and integration on manifolds

Theorem. Let $M \subset \mathbb{R}^n$ be a C^k m-manifold. Prove that M admits an atlas $\{(U_\alpha, B_\alpha, \phi_\alpha)\}_{\alpha \in A}$ satisfying the following properties.

1. A is countable, and if M is compact then A is finite.
2. For each $\alpha \in A$ the set $B_\alpha \subset \mathbb{R}^m$ is an open ball.
3. For each $\alpha \in A$ there exists an open set $V_\alpha \subset \mathbb{R}^m$ such that $\varnothing \neq B_\alpha \subset \overline{B_\alpha} \subset V_\alpha$ with $\overline{B_\alpha}$ compact, and $\phi_\alpha \in C^k(B_\alpha; \mathbb{R}^n)$ is the restriction to B_α of a map $\psi_\alpha \in C^k(V_\alpha; \mathbb{R}^n)$ such that $\psi_\alpha(V_\alpha) \subset M$. In particular ϕ_α and all of its derivatives up to order k are bounded on B_α.
4. $M = \bigcup_{\alpha \in A} \phi_\alpha(B_\alpha) = \bigcup_{\alpha \in A} \phi_\alpha(\overline{B_\alpha})$.
5. $0 < \mathcal{H}^m(\phi_\alpha(\overline{B_\alpha})) < \infty$ for every $\alpha \in A$.

Proof. For each $z \in M$, there exists C^k local coordinates (W_z, V_z, ϕ_z). Let $x \in \phi_z^{-1}(z) \in V_z$ and pick a ball B_z such that $x \in B_z \subset \overline{B_z} \subset V_z$. Let ϕ_z be the restriction on B_z. Note that $\phi_z(B_z)$ is relatively open, so we can select open set $U_z \subset \mathbb{R}^n$ such that $U_z \cap M = \phi_z(B_z)$. It is easy to check that (U_z, B_z, ϕ_z) is C^k local coordinates at $z \in M$, and thus $\{(U_z, B_z, \phi_z)\}_{z \in M}$ is an atlas of M.

Now $\{U_z\}_{z \in M}$ is a open cover for M. If M is compact, select a finite subcover. Otherwise, select a countable subcover by selecting a U_z for each rational ball with rational radius (Linderlöf's theorem). Denote the corresponding atlas by $\{(U_\alpha, B_\alpha, \phi_\alpha)\}_{\alpha \in A}$. We proceed to show it satisfies the desired properties.

1. $\mathcal{A}$ is clearly countable. Now suppose M is compact, then $\{U_\alpha\}_{\alpha \in A}$ admits a finite subcover. By construction, $\mathcal{A}$ is then finite.
2. For each $\alpha \in A$, B_α is constrcted to be an open ball in $\mathbb{R}^m$.
3. It is clear that $\overline{B_\alpha}$ is compact. Moreover, by construction $B_\alpha \subset \overline{B_\alpha} \subset \widetilde{V}_{\lambda_i}$, and ϕ_α is the restriction to B_α of the map $\widetilde{\phi}_{\lambda_i}$, and $\widetilde{\phi}_{\lambda_i}(\widetilde{V}_{\lambda_i}) \subset M$ because $\lambda_i \in \Lambda$ is taken from the original atlas $\{(U_\lambda, V_\lambda, \phi_\lambda)\}_{\lambda \in \Lambda}$.

4. We already know that

$$U_\alpha \cap M = \phi_\alpha(B_\alpha) \subset \phi_\alpha(\overline{B_\alpha}) \subset \widetilde{\phi}_{\lambda_i}(\widetilde{V}_{\lambda_i}) \subset M.$$

   We also know that $\{U_\alpha\}_\alpha$ covers M. Therefore,

$$M = \bigcup_{\alpha \in A} \phi_\alpha(B_\alpha) = \bigcup_{\alpha \in A} \phi_\alpha(\overline{B_\alpha})$$

5. Note that $\overline{B_\alpha}$ is compact and ϕ_α and in particular all its derivatives up to order k are bounded on $\overline{B_\alpha}$. It follows that $\mathcal{H}^m(\overline{B_\alpha}) < \infty$ and ϕ_α is Lipshitz on $\overline{B_\alpha}$. Therefore, $\mathcal{H}^m(\phi_\alpha(\overline{B_\alpha})) < \infty$.

   Now, to show $\mathcal{H}^m(\phi(\overline{B_\alpha})) > 0$, note that ϕ_α is injective. Then, from the area formula we have

$$\mathcal{H}^m(\phi_\alpha(\overline{B_\alpha})) = \int_{\overline{B_\alpha}} J_{\phi_\alpha}\, d\lambda$$

   However, for all $x \in \overline{B_\alpha}$, $D\phi_\alpha(x)$ has rank m so $D\phi_\alpha(x)$ is injective. This implies that $J_{\phi_\alpha}(x) > 0$ for all $x \in \overline{B_\alpha}$. Note also $\overline{B_\alpha}$ is compact so there exists $\varepsilon > 0$ such that $J_{\phi_\alpha} > \varepsilon$ on $\overline{B_\alpha}$. Therefore,

$$\mathcal{H}^m(\phi_\alpha(\overline{B_\alpha})) \geq \varepsilon \mathcal{H}^m(\overline{B_\alpha}) > 0,$$

   as desired.

□

Theorem. Let $M \subset \mathbb{R}^n$ be a C^1 m-manifold. Prove that the following hold.

1. $M = \bigcup_{k=0}^{\infty} K_k$, where $K_k \subset \mathbb{R}^n$ is compact and $\mathcal{H}^m(K_k) < \infty$. In particular, M is σ-compact.
2. For each $E \subset M$ there exists $G \in G_\delta(M)$ such that $\mathcal{H}^m(E) = \mathcal{H}^m(G)$, and in particular the outer measure $\mathcal{H}^m$ is cover-regular on M.
3. $(M, \mathfrak{H}^m(M), \mathcal{H}^m)$ is a complete and σ-finite measure space such that $\mathfrak{B}(M) \subset \mathfrak{H}^m(M)$. Moreover, if $E \subset M$ satisfies $\mathcal{H}^m(E) = 0$ then $E \in \mathfrak{H}^m(M)$.
4. $\mathfrak{H}^m(M) = \mathfrak{H}^m(\mathbb{R}^n)_M$.
5. The following are equivalent for $E \subset M$.
   (a) E is measurable with respect to the outer measure $\mathcal{H}^m$, i.e. $E \in \mathfrak{H}^m(M)$.
   (b) There exists $G \in G_\delta(M)$ such that $E \subset G$ and $\mathcal{H}^m(G \setminus E) = 0$.
   (c) There exists $F \in F_\sigma(M)$ such that $F \subset E$ and $\mathcal{H}^m(E \setminus F) = 0$.
   (d) For every $\varepsilon > 0$ there exists an open set $U \subset M$ such that $E \subset U$ and $\mathcal{H}^m(U \setminus E) < \varepsilon$.
   (e) For every $\varepsilon > 0$ there exists a closed set $C \subset M$ such that $C \subset E$ and $\mathcal{H}^m(E \setminus C) < \varepsilon$.

6. We have that

$$\mathfrak{H}^m(M) = \{E \cup F : E \in \mathfrak{B}(M) \text{ and } F \subset N \text{ for } N \in \mathfrak{B}(M) \text{ such that } \mathcal{H}^m(N) = 0\}.$$

7. $\dim_H(M) = m$.

Proof. Let $\{(U_\alpha, B_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas for M that satisfies the properties listed in the previous theorem.

1. Note that $M = \bigcup_{\alpha \in A} \phi_\alpha(\overline{B_\alpha})$. Moreover, we know $\overline{B_\alpha}$ is compact so $\phi_\alpha(\overline{B_\alpha})$ is compact, and $\mathcal{H}^m(\phi_\alpha(\overline{B_\alpha})) < \infty$ for all $\alpha \in A$.
2. Now M is a metric space and $M = \bigcup_{k=0}^\infty K_k$ where $K_k \subset \mathbb{R}^n$ is compact with $\mathcal{H}^m(K_k) < \infty$. Therefore by a previous theorem, for each $E \subset M$, there exists $G \in G_\delta(M)$ such that $\mathcal{H}^m(E) = \mathcal{H}^m(G)$. In particular, the outer measure $\mathcal{H}^m$ is cover-regular on M.
3. By a previous theorem, $(M, \mathfrak{H}^m(M), \mathcal{H}^m)$ is a complete and σ-finite measure space such that $\mathfrak{B}(M) \subset \mathfrak{H}^m(M)$, and $\mathcal{H}^m(E) = 0$ implies $E \in \mathfrak{H}^m(M)$.
4. Note that $M = \bigcup_{k=0}^\infty K_k$ where $\mathcal{H}^m(K_k) < \infty$, so a previous theorem gives $\mathfrak{H}^m(M) = \mathfrak{H}^m(\mathbb{R}^n)_M$.
5. Note that $M = \bigcup_{\alpha \in A} \phi_\alpha(B_\alpha)$, $\phi_\alpha(B_\alpha) \subset M$ relatively open in M, and

$$\mathcal{H}^m(\phi_\alpha(B_\alpha)) \leq \mathcal{H}^m \phi_\alpha(\overline{B_\alpha}) < \infty.$$

   Therefore by a previous theorem, the items (a) to (e) are equivalent.
6. A previous theorem on Hausdorff measure space gives that

$$\mathfrak{H}^m(M) = \{E \cup F : E \in \mathfrak{B}(M) \text{ and } F \subset N \text{ for } N \in \mathfrak{B}(M) \text{ such that } \mathcal{H}^m(N) = 0\}.$$

7. Note that $M = \bigcup_{\alpha \in A} \phi_\alpha(\overline{B_\alpha})$ and $0 < \mathcal{H}^m(\phi_\alpha(\overline{B_\alpha})) < \infty$. This implies that $\mathcal{H}^m(M) \leq \infty$ so $\dim_H(M) \leq m$. On the other hand, for $0 \leq s < m$, we have $\mathcal{H}^s(\phi_\alpha(\overline{B_\alpha})) = 0$. It follows that

$$\mathcal{H}^s(M) = \sum_{\alpha \in A} \mathcal{H}^s(\phi_\alpha(\overline{B_\alpha})) = 0.$$

   Therefore, $\dim_H(M) \geq m$. This implies that $\dim_H(M) = m$.

□

4.1.3 Tangent spaces

Lemma. Suppose (U_0, V_0, ϕ_0) and (U_1, V_1, ϕ_1) are C^k local m-coordinates at $z \in M$. Suppose also $U_0 \cap U_1 \neq \emptyset$ and $z \in U_0 \cap U_1 \cap M$. Then the following holds:

1. $D\phi_0(\phi_0^{-1}(z)) = D\phi_1(\phi_1^{-1}(z)) D[\phi_1^{-1} \circ \phi_0](\phi_0^{-1}(z))$.

2. $\operatorname{range} D\phi_0(\phi_0^{-1}(z)) = \operatorname{range} D\phi_1(\phi_1^{-1}(z))$

Proof. Let $a = \phi_0^{-1}(z)$ and $b = \phi_1^{-1}(z)$, then

$$\phi_0 = \phi_1 \circ \phi_1^{-1} \circ \phi_0 = \phi_1 \circ \zeta.$$

It follows that $D\phi_0 = [D\phi_1 \circ \zeta] \circ D\zeta$, and

$$D\phi_0(a) = [D\phi_1 \circ \zeta(a)]D\zeta(a) = [D\phi_1(b)]D\zeta(a).$$

This proves item 1. However, ζ is a diffeomorphism, so

$$\operatorname{range} D\phi_0(\phi_0^{-1}(z)) = \operatorname{range} D\phi_1(\phi_1^{-1}(z)).$$

□

Definition (tangent space)**.** Define the **tangent space** to M at z to be $T_zM = \operatorname{range} D\phi(\phi^{-1}(z))$ for **any** C^k m-local coordinate chart (U, V, ϕ).

Example. Let (U, V, ϕ) be m-coordinates at $z \in M$ and let $x = \phi^{-1}(z)$. Then,

$$T_zM = \operatorname{range} D\phi(x) = \operatorname{span}\{D\phi(x)e_1, \dots, D\phi(x)e_m\} = \operatorname{span}\{\partial_1\phi(x), \dots, \partial_m\phi(x)\}.$$

△

Theorem. Let $\emptyset \neq U \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^r$ be C^k for $1 \le r < n$. We saw before that if $z_0 \in U$, and

$$M = \{z \in U : f(z) = f(z_0) \text{ and } Df(z) \text{ has rank } r\} \neq \emptyset,$$

then M is a C^k $(n-r)$-manifold. If $z \in M$, then $T_zM = \ker Df(z)$.

Proof. Let $(\widetilde{U}, V, \phi)$ be local coordinate at $z \in M$. Then $f(\phi(x)) = f(z_0)$ for all $x \in C$. Then,

$$0 = Df(\phi(x))D\phi(x).$$

Therefore, for $x_0 \in V$ such that $\phi(x_0) = z$, we have $0 = Df(z)D\phi(x_0)$. This implies that $T_zM = \operatorname{range} D\phi(x_0) \subset \ker Df(z)$. However,

$$\dim \operatorname{range} D\phi(x_0) = \dim \ker Df(z) = n - r,$$

so $T_zM = \ker Df(z)$. □

Example. 1. Let $M = \{x \in \mathbb{R}^n : Sx \cdot x = 1\}$ for $S \in \mathbb{R}^{n\times n}_{\text{sym}}$ with at least one positive eigenvalue. Define $f(x) = Sx \cdot x$, then $Df(x)h = 2Sx \cdot h$. This then implies that

$$T_xM = \{h : Sx \cdot h = 0\}.$$

2. Let $\Gamma(f) = \{(x, f(x)) : x \in U\}$ for $f : U \to \mathbb{R}^m$ and $U \subset \mathbb{R}^n$ open. Define $F(x, y) = f(x) - y$, then $DF(x, y)(h, k) = Df(x)h - k$. This then implies that

$$T_{(x,f(x))}\Gamma(f) = \{(h, Df(x)h) : h \in \mathbb{R}^n\}.$$

3. Recall that $\mathrm{SL}(n) = \{M \in \mathbb{R}^{n\times n} : \det M = 1\}$. Define $f(x) = \det M$, then $Df(M)N = \det M \operatorname{tr}(M^{-1}N)$. This then implies that

$$T_M \mathrm{SL}(n) = M\mathbb{R}^{n\times n}_{\text{tr-0}},$$

where $\mathbb{R}^{n\times n}_{\text{tr-0}} = \{N \in \mathbb{R}^{n\times n} : \operatorname{tr} N = 0\}$

4. Let $O(n) = \{M \in \mathbb{R}^{n\times n} : M^T M = I\}$. Define $f(M) = M^T M - I$, then $Df(M)N = M^T N + N^T M$. This then implies that

$$T_M O(n) = M\mathbb{R}^{n\times n}_{\text{asym}}.$$

$\triangle$

4.1.4 Mappings between manifolds

Lemma. Suppose $M \subset \mathbb{R}^{n_1}$ and $N \subset \mathbb{R}^{n_2}$ are C^k m_1-manifold and m_2-manifold respectively. Suppose also that $f : M \to N$. Let x be such that $\phi_1(x) = z \in U_0 \cap U_1 \cap M$ and $f(z) \in \widetilde{U_0} \cap \widetilde{U_1} \cap N$. Then the following holds:

1. $\psi_0^{-1} \circ f \circ \phi_0$ is differentiable at x if and only if $\psi_1^{-1} \circ f \circ \phi_1$ is differentiable at $\zeta(x) = \phi_1^{-1} \circ \phi_0^{-1}(x)$. For $1 \le j \le k$, $\psi_0^{-1} \circ f \circ \phi_0$ is C^j at x if and only if $\psi_1^{-1} \circ f \circ \phi_1$ is C^j at $\zeta(x) = \phi_1^{-1} \circ \phi_0^{-1}(x)$.
2. Given $v \in T_z M = \operatorname{range} D\phi_i(\phi_i^{-1}(z))$, there exists a unique $w \in T_{f(z)}N$ such that if for $i = 0, 1$ we have $v = D\phi_i(\phi_i^{-1})u_i$, then

$$w = D\psi_i(\psi_i^{-1}(f(z)))D[\psi_i^{-1} \circ f \circ \phi_i](\phi_i^{-1}(z))u_i$$

for $i = 0, 1$.

Proof. ***** TO-DO ***** $\square$

Definition. Suppose $M_1 \subset \mathbb{R}^{n_1}$ and $M_2 \subset \mathbb{R}^{n_2}$ are C^k m_1-manifold and m_2-manifold respectively. Suppose also that $f : M_1 \to M_2$.

1. We say f is differentiable at $z \in M$ if $\psi^{-1} \circ f \circ \phi$ is differentiable at $\phi^{-1}(z)$ for any choice of C^k local coordinates at z and $f(z)$. We define $Df(z) \in \mathcal{L}(T_z M_1, T_{f(z)} M_2)$ via $v \mapsto w$ from the previous lemma.
2. We say $f \in C^j(M_1, M_2)$ for $1 \le j \le k$ if $\psi^{-1} \circ f \circ \phi$ is C^j.

The previous lemma ensures that these concepts are well-defined.

Theorem (Chain rule)**.** Suppose M_1, M_2, M_3 are nonempty C^k manifolds. Suppose $f \in C^k(M_1; M_2)$ and $g \in C^k(M_2; M_3)$. Then $g \circ f \in C^k(M_1; M_3)$ and for each $z \in M_1$,

$$D(g \circ f)(z) = Dg(f(z))Df(z).$$

Proof. ***** TO-DO ***** $\square$

Definition. Let M_1 and M_2 be C^1 manifolds and $f : M_1 \to M_2$ be C^1.

1. We say f is a **submersion** at $z \in M$ if $Df(z) \in \mathcal{L}(T_z M_1; T_{f(z)} M_2)$ is surjective. We say f is a submersion if f is a submersion at each $z \in M$.
2. We say f is an **immersion** at $z \in M$ if $Df(z) \in \mathcal{L}(T_z M_1; T_{f(z)} M_2)$ is injective. We say f is an immersion if f is an immersion at each $z \in M$.
3. We say f is a **embedding** if it is an immersion that is a homeomorphism from M_1 to $f(M_1) \subset M_2$.

Definition (diffeomorphism)**.** Let M_1 and M_2 be C^k manifolds. We say $f \in C^k(M_1; M_2)$ is a **diffeomorphism** if f is a homeomorphism and $f^{-1} \in C^k(M_2; M_1)$.

Proposition. The following holds for manifolds M_1, M_2:

1. If $f : M_1 \to M_2$ is a diffeomorphism, then f, f^{-1} are embeddings.
2. If M_1 and M_2 are diffeomorphic, then M_1 and M_2 have the same dimension.

Proof. From the chain rule, we know $Df^{-1}(f(z)) = [Df(z)]^{-1}$. Therefore, $T_z M_1$ and $T_{f(z)} M_2$ are isomophic for each $z \in M$. This also shows that f, f^{-1} are immersions. $\square$

4.1.5 Orientation on manifolds

***** TO-DO *****

4.1.6 Unit normals and orientations on codimension one manifolds

***** TO-DO *****

4.1.7 Manifolds with boundary

Definition. Let $z \in M \subset \mathbb{R}^n$. WE say z admits C^k **local boundary coordinates** if

1. there exists open set $U \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ such that $z \in U$ and $W \cap \{x \in \mathbb{R}^m : x_m = 0\} \neq \emptyset$.
2. There exists homeomorphism $\phi : V \to \phi(V) = U \cap M$ where $V = W \cap \{x \in \mathbb{R}^m : x_m \geq 0\}$.
3. $\phi = \widetilde{\phi}|_V$ where $\widetilde{\phi} \in C^k(W; \mathbb{R}^n)$ is such that $D\widetilde{\phi}$ has rank m everywhere.

Proposition. Given a point $z \in M \subset \mathbb{R}^n$, **at most one** of the following can be true:

1. z admits C^k local m-coordinates.

2. z admits C^k local boundary coordinates.

Proof. ***** TO-DO ***** $\square$

Definition. A C^k m-manifold with boundary is a set $\emptyset \neq M \subset \mathbb{R}^n$ such that $M = M^\circ \cup \partial M$, where $M^\circ \neq \emptyset$, $\partial M \neq \emptyset$, and

$$\begin{aligned} M^\circ &= \{z \in M : z \text{ admits local } C^k \text{ } m\text{-coordinates}\}, \\ \partial M &= \{z \in M : z \text{ admits local } C^k \text{ boundary coordinates}\}. \end{aligned}$$

We call M° the **interior** of M and ∂M the **boundary** of M. Note that $M^\circ \cap \partial M = \emptyset$.

We have seen before that level sets are manifolds. Here we have another version for manifolds with boundary.

Theorem (superlevel sets as manifolds with boundary)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be open and suppose that $f \in C^k(U; \mathbb{R}^r)$ for $1 \leq r < n$. If $r \geq 2$, write $f = (\widetilde{f}, f_r)$, i.e. $\widetilde{f} = \sum_{i=1}^{r-1} f_i e_i$. Let $z_0 \in U$ and set

$$L = \{z \in U : f(z) = f(z_0)\}$$

and

$$S = \{z \in U : \widetilde{f}(z) = \widetilde{f}(z_0) \text{ and } f_r(z) \geq f_r(z_0)\}$$

with the understanding that the $\widetilde{f}$ condition is ignored if $r = 1$. Finally, suppose that $Df(z)$ has rank r for every $z \in L$ and $D\widetilde{f}(z)$ has rank $r - 1$ for every $z \in S$. Prove that if S and L are nonempty, then S is a C^k $(n - r + 1)$-manifold with boundary and that $\partial S = L$.

Proof. Suppose S and L are nonempty.

We know that

$$\widetilde{S} = \{z \in U : \widetilde{f}(z) = \widetilde{f}(z_0) \text{ and } D\widetilde{f}(z) \text{ has rank } r - 1\}$$

is a C^k $(n - r + 1)$-manifold. This implies that for each $z \in \widetilde{S}$, there exists open set $W \subset \mathbb{R}^n$ with $z \in U$, open set $V \subset \mathbb{R}^{n-r+1}$ and homeomorphism $\phi : V \to \phi(V) = W \cap \widetilde{S}$ such that ϕ is C^k and $D\phi(x)$ has rank $n - r + 1$ for all $x \in W$. Now, for $z \in S$ is such that $f_r(z) > f_r(z_0)$ and let (W, V, ϕ) be the local coordinates correspond to z in $\widetilde{S}$. Then, because ϕ is a homeomorphism, we can shrink W and V such that $f_r(z) > f_r(z_0)$ for all $z \in W$. This then implies that $\phi(V) = W \cap S$ and the new (W, V, ϕ) is a local coordinates for $z \in S$.

It remains to find local boundary coordinates for $z \in L$. Since $Df(z)$ has rank r, $Df(z)$ is surjective. The right inverse function theorem then provides open sets $W \subset \mathbb{R}^r$ with $f(z_0) \in W$ and $V \subset \mathbb{R}^{n-r}$ and a C^k diffeomorphism $F : W \times V \to F(W \times V) \subset U$ such that $f(F(w, v)) = w$ for all $(w, v) \in W \times V$. Let $\varepsilon > 0$ be small enough such that

$$\{\widetilde{f}(z_0)\} \times (f_r(z_0) - \varepsilon, f_r(z_0) + \varepsilon) \subset W.$$

Now we define a function $\phi : V \times (-\varepsilon, \varepsilon)$, via

$$\phi(v, c) = F(\widetilde{f}(z_0), f_r(z_0) + c, v).$$

It is easy to verify that this gives a local boundary coordinates at $z \in L$. Indeed, F being a C^k diffeomorphism guarantees that $D\phi(v, c)$ has rank $n - r + 1$ for all $(v, c) \in V \times (-\varepsilon, \varepsilon)$. We can now conclude that S is a C^k $(n - r + 1)$-manifold with boundary and $\partial S = L$. ☐

4.2 Differential forms

We can integrate functions on manifolds with respect to $\mathcal{H}^m$, but manifolds come equipped with tangent spaces, which these function cannot easily interact with. Therefore, we want some sort of mathematical structure that

1. interacts easily with $T_z M$ at each $z \in M$,
2. "be integrable",
3. interact nicely with geometry.

The solution is as follows:

1. We want a map f from the domain M and we hope $f(z)$ to interact with $T_z M$ is a simple manner. Therefore, we require $f : M \to \mathcal{L}^k(\mathbb{R}^n; \mathbb{R})$, which we call a "tensor field".
2. We want $f(z) \in \mathcal{L}^k(\mathbb{R}^n; \mathbb{R})$ to interact with the "genuine k-dimensional structure", in the way that $f(z)(v_1, \ldots, v_k) = 0$ if $\{v_1, \ldots, v_k\}$ are linearly dependent. However, if $\omega \in \mathcal{L}^k(V; R)$, then $\omega(v_1, \ldots, v_k) = 0$ for all linearly dependent $\{v_1, \ldots, v_k\}$ if and only if ω is anti-symmetric.

Combining these requirements, we want to study $f : M \to \Lambda^k(\mathbb{R}^n)$, where

$$\Lambda^k(\mathbb{R}^n) = \{\omega \in \mathcal{L}^k(\mathbb{R}^n; \mathbb{R}) : \omega \text{ is anti-symmetric}\}$$

is the set of k-forms.

Recall the definition and properties of antisymmetric multilinear maps:

Definition. Let $k \in \mathbb{N}^+$. Let V, W be vector spaces over $\mathbb{F}$.

1. $T \in L^k(V, W)$ is said to be **symmetric** if for every $v_1, \ldots, v_k \in V$ and $P \in S_k$, we have
$$T(v_1, \ldots, v_k) = T(v_{P(1)}, \ldots, v_{P(k)}).$$
2. $T \in L^k(V, W)$ is said to be **antisymmetric** if for every $v_1, \ldots, v_k \in V$ and $P \in S_k$, we have
$$T(v_{P(1)}, \ldots, v_{P(K)}) = \operatorname{sgn}(P) T(v_1, \ldots, v_k).$$

Theorem. Let $k \in \mathbb{N}^+$ and V, W be vector spaces over $\mathbb{F}$. Suppose $T \in L^k(V, W)$. Then following are equivalent:

1. T is antisymmetric.
2. For every $v_1, \dots, v_k \in V$ and transposition $P \in S_k$, we have

$$T(v_1, \dots, v_k) = -T(v_{P(1)}, \dots, v_{P(k)}).$$

3. If $v_1, \dots, v_k$ is such that $v_i = v_j$ for some $i \neq j$, then

$$T(v_1, \dots, v_k) = 0.$$

4. If $v_1, \dots, v_l \in V$ are linearly dependent, then

$$T(v_1, \dots, v_k) = 0.$$

Some algebra facts follows from the definition.

1. If $\omega \in \Lambda^k$ and $\eta \in \Lambda^j$, then

$$\omega \wedge \eta = \frac{(k+j)!}{k!\, j!} \mathcal{A}(\omega \otimes \eta) \in \Lambda^{k+j}(\mathbb{R}^n),$$

where $\mathcal{A}$ represents the anti-symmetrization given by

$$\mathcal{A}T(v_1, \dots, v_k) = \frac{1}{k!} \sum_{P \in S_k} \operatorname{sgn}(P) T(v_1, \dots, v_k)$$

for any $T \in L^k(\mathbb{R}^n)$. However, this is not qiute commutative, as

$$\omega \wedge \eta = (-1)^{jk} \eta \wedge \omega.$$

but it is associative:

$$\omega \wedge (\eta \wedge \theta) = (\omega \wedge \eta) \wedge \theta = \omega \wedge \eta \wedge \theta.$$

2. We have

$$\dim \Lambda^k(\mathbb{R}^n) = \begin{cases} \binom{n}{k} & \text{if } 0 \leq k \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

A basis of $\Lambda^k(\mathbb{R}^n)$ is $\{e^{i_1} \wedge \cdots \wedge e^{i_k} : i_1 < \cdots < i_k\}$ for $e^i(x) = x_i = e_i \cdot x$. Recall that $\mathcal{A}(n,k) = \{\alpha = (\alpha_1, \dots, \alpha_k) : 1 \leq \alpha_1 < \cdots < \alpha_k \leq n\}$. Write

$$e^\alpha = e^{\alpha_1} \wedge \cdots \wedge e^{\alpha_k}.$$

We also set $\mathcal{A}(n,k) = \emptyset$ for $k \geq n+1$. Therefore, any $\omega \in \Lambda^k(\mathbb{R}^n)$ can be written as

$$\omega = \sum_{\alpha \in \mathcal{A}(n,k)} \omega_\alpha e^\alpha.$$

Proposition. Suppose $\omega_1, \dots, \omega_k \in \Lambda^1(\mathbb{R}^n)$, then

$$(\omega_1 \wedge \dots \wedge \omega_k)(v_1, \dots, v_k) = \det V,$$

where $V \in \mathbb{R}^{k\times k}$ is such that $V_{ij} = \omega_i(v_j)$.

Proof. *** **TO-DO** *** □

Corollary. Suppose $1 \le k \le n$ and $\alpha \in \mathcal{A}(n,k)$. If $v_1, \dots, v_k \in \mathbb{R}^n$ then

$$e^\alpha(v_1, \dots, v_k) = \det U,$$

where $U \in \mathbb{R}^{k\times k}$ is such that $U_{ij} = v_j \cdot e_{\alpha_i}$.

Corollary. If $\omega \in \Lambda^k(\mathbb{R}^n)$, then $\omega \in \sum_{\alpha\in\mathcal{A}(n,k)} \omega_\alpha e^\alpha$, where

$$\omega_\alpha = \omega(e_{\alpha_1}, \dots, e_{\alpha_k}).$$

Proposition. Suppose $\omega \in \Lambda^k(\mathbb{R}^n)$, then

$$\omega(v_1, \dots, v_n) = \omega(e_1, \dots, e_n) \det V,$$

where $V \in \mathbb{R}^{n\times n}$ is such that $V_{ij} = v_j \cdot e_i$.

Proof. *** **TO-DO** *** □

Definition (exterior derivative)**.** Let $\emptyset \ne U \subset \mathbb{R}^n$ open and $1 \le r \le \infty$.

1. Given $\omega \in \Omega^{0,r}(U)$, we set

$$d\omega = \sum_{i=1}^n \partial_i \omega e^i \in \Omega^{1,r-1}(U).$$

2. Given $\omega \in \Omega^{k,r}(U)$, write $\omega = \sum_{\alpha\in\mathcal{A}(n,k)} \omega_\alpha e^\alpha$. We then define

$$d\omega = \sum_{\alpha\in\mathcal{A}(n,k)} d\omega_\alpha \wedge e^\alpha = \sum_{\alpha\in\mathcal{A}(n,k)} \sum_{i=1}^n \partial_i \omega_\alpha e^i \wedge e^\alpha \in \Omega^{k+1,r-1}(U).$$

We call $d\omega$ the **exterior derivative** of ω.

Theorem. Let $\emptyset \ne U \subset \mathbb{R}^n$ be open and $r \ge 1$. Then the following holds:

1. $d : \Omega^{k,r}(U) \to \Omega^{k+1,r-1}(U)$ is linear.
2. If $\omega \in \Omega^{k,r}(U)$ and $\zeta \in \Omega^{j,r}(U)$, then

$$d(\omega \wedge \zeta) = d\omega \wedge \zeta + (-1)^k \omega \wedge d\zeta.$$

3. **Suppose** $r \geq 2$**, then** $d^2 = d \circ d = 0$**.**

Proof. *** **TO-DO** *** □

Definition (pullback)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ and $\emptyset \neq V \subset \mathbb{R}^m$ be open and $\phi : V \to U$ be differentiable. For $\omega \in \Omega^k(U)$, we define $\phi^*\omega \in \Omega^k(V)$, the **pullback** if ω by ϕ, by

$$\phi^*\omega(v_1, \ldots, v_k) = [\omega \circ \phi(x)](D\phi(x)v_1, \ldots, D\phi(x)v_k),$$

for $v_1, \ldots, v_k \in \mathbb{R}^m$ and $x \in V$.

Theorem. Suppose $\emptyset \neq U \subset \mathbb{R}^n$ and $\emptyset \neq V \subset \mathbb{R}^m$ open, and $\phi : V \to U$ be differentiable. The following holds:

1. If $\phi \in C^r$ with $r \geq 1$, then the map $\phi^* : \Omega^{k,r-1}(U) \to \Omega^{k,r-1}(V)$ is linear and well-defined.
2. $\phi^* e^i = \sum_{j=1}^m \partial_j \phi_i e^j = d\phi_i$ for each $1 \leq i \leq n$.
3. If $\omega \in \Omega^k(U)$ and $\zeta \in \Omega^j(U)$, then
$$\phi^*(\omega \wedge \zeta) = (\phi^*\omega) \wedge (\phi^*\zeta).$$
4. If $\omega \in \Omega^{k,1}(U)$ and $\phi \in C^r$ **with** $r \geq 2$, then
$$\phi^*(d\omega) = d(\phi^*\omega).$$
5. Let $\omega = \sum_{\alpha \in \mathcal{A}(n,k)} \omega_\alpha e^\alpha$, then
$$\begin{aligned} \phi^*\omega &= \sum_{\alpha \in \mathcal{A}(n,k)} \omega_\alpha \circ \phi d\phi_1 \wedge \cdots \wedge d\phi_k \\ &= \sum_{\beta \in \mathcal{A}(m,k)} \left[\sum_{\alpha \in \mathcal{A}(n,k)} \omega_\beta \circ \phi \det D_\beta \phi_\alpha \right] e^\beta, \end{aligned}$$
where $D_\beta \phi_\alpha \in \mathbb{R}^{k \times k}$ is such that $(D_\beta \phi_\alpha)_{ij} = \partial_{\beta_j} \phi_{\alpha_i}$.

Proof. *** **TO-DO** *** □

4.3 Stokes-Cartan and fundamental theorems of vector calculus

4.3.1 Integration of differential forms on orientated manifolds

Lemma. Suppose that $V \subset \mathbb{R}^n$ is a subspace of dimension $1 \leq m \leq n$. Given $v_1, \ldots, v_m \in V$ write $\langle v_1, \ldots, v_m \rangle \in \mathbb{R}^{n \times m}$ for the matrix with j-th column given by v_j for $1 \leq j \leq m$. Suppose that $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_m)$ are ordered bases of V, and let $B \in \mathbb{R}^{m \times m}$ denote the change of basis matrix from $(v_1, \ldots, v_m)$ to $(w_1, \ldots, w_m)$, i.e. $v_i = \sum_{j=1}^m B_{ij} w_j$ for each $1 \leq i \leq m$.

1. Let $\alpha \in \mathcal{A}(n,m)$. We then have
$$\frac{\det\langle v_1,\dots,v_m\rangle_\alpha}{[\![\langle v_1,\dots,v_m\rangle]\!]} = \frac{\det B}{|\det B|}\frac{\det\langle w_1,\dots,w_m\rangle_\alpha}{[\![\langle w_1,\dots,w_m\rangle]\!]},$$
and in particular if $\det B > 0$ then
$$\frac{\det\langle v_1,\dots,v_m\rangle_\alpha}{[\![\langle v_1,\dots,v_m\rangle]\!]} = \frac{\det\langle w_1,\dots,w_m\rangle_\alpha}{[\![\langle w_1,\dots,w_m\rangle]\!]}.$$

2. Suppose now that $\omega \in \mathcal{L}^m(\mathbb{R}^n;\mathbb{R})$ is antisymmetric. We then have
$$\frac{\omega(v_1,\dots,v_m)}{[\![\langle v_1,\dots,v_m\rangle]\!]} = \frac{\det B}{|\det B|}\frac{\omega(w_1,\dots,w_m)}{[\![\langle w_1,\dots,w_m\rangle]\!]},$$
and in particular if $\det B > 0$ then
$$\frac{\omega(v_1,\dots,v_m)}{[\![\langle v_1,\dots,v_m\rangle]\!]} = \frac{\omega(w_1,\dots,w_m)}{[\![\langle w_1,\dots,w_m\rangle]\!]}.$$

Proof. 1. Let $V = \langle v_1,\dots,v_m\rangle$ and $W = \langle w_1,\dots,w_m\rangle$. Then we have
$$V_{ki} = (v_i)_k = B_{ij}(w_j)_k = B_{ij}W_{kj}$$
It follows that $V = WB^T$. Moreover, for all $\alpha \in \mathcal{A}(n,m)$, we have $V_\alpha = W_\alpha B^T$. Therefore,
$$\begin{aligned}\frac{\det\langle v_1,\dots,v_m\rangle_\alpha}{[\![\langle v_1,\dots,v_m\rangle]\!]} &= \frac{\det V_\alpha}{\sqrt{\det V^TV}} = \frac{\det W_\alpha B^T}{\sqrt{\det BW^TWB^T}}\\ &= \frac{\det B}{|\det B|}\frac{\det W_\alpha}{\sqrt{\det W^TW}} = \frac{\det B}{|\det B|}\frac{\det\langle w_1,\dots,w_m\rangle_\alpha}{[\![\langle w_1,\dots,w_m\rangle]\!]}.\end{aligned}$$
In particular, $\det B > 0$, then $\frac{\det B}{|\det B|} = 1$ and
$$\frac{\det\langle v_1,\dots,v_m\rangle_\alpha}{[\![\langle v_1,\dots,v_m\rangle]\!]} = \frac{\det\langle w_1,\dots,w_m\rangle_\alpha}{[\![\langle w_1,\dots,w_m\rangle]\!]}.$$

2. Again we have
$$[\![\langle v_1,\dots,v_m\rangle]\!] = |\det B|\,[\![\langle w_1,\dots,w_m\rangle]\!].$$
For the numerator, we have
$$\begin{aligned}\omega(v_1,\dots,v_m) &= \omega\left(\sum_{j=1}^m B_{1j}w_j,\ \cdots,\ \sum_{j=1}^m B_{mj}w_j\right)\\ &= \sum_{j_1=1}^m\cdots\sum_{j_m=1}^m\prod_{k=1}^m B_{kj_k}\omega(w_{j_1},\cdots,w_{j_m}).\end{aligned}$$

However, we know $\omega(w_{j_1},\dots,w_{j_m})=0$ if there exists $k\neq l$ such that $w_{j_k}=w_{j_l}$. It follows that

$$\begin{aligned}\omega(v_1,\dots,v_m)&=\sum_{P\in S_m}\prod_{k=1}^m B_{kP(k)}\omega(w_{P(1)},\dots,w_{P(m)})\\&=\sum_{P\in S_m}\operatorname{sgn}(P)\prod_{k=1}^m B_{kP(k)}\omega(w_1,\dots,w_m)\\&=\omega(w_1,\dots,w_m)\det B,\end{aligned}$$

where we used the fact that

$$\det B=\sum_{P\in S_m}\operatorname{sgn}(P)\prod_{k=1}^m B_{kP(k)}.$$

This implies that

$$\frac{\omega(v_1,\dots,v_m)}{[\![\langle v_1,\dots,v_m\rangle]\!]}=\frac{\det B}{|\det B|}\frac{\omega(w_1,\dots,w_m)}{[\![\langle w_1,\dots,w_m\rangle]\!]}.$$

In particular, if $\det B>0$, then $\frac{\det B}{|\det B|}=1$, and

$$\frac{\omega(v_1,\dots,v_m)}{[\![\langle v_1,\dots,v_m\rangle]\!]}=\frac{\omega(w_1,\dots,w_m)}{[\![\langle w_1,\dots,w_m\rangle]\!]},$$

as desired.

□

Lemma. Let $\omega\in\Lambda^m(\mathbb{R}^n)$ and consider $V\subset\mathbb{R}^n$ a subspace of dimension $m\leq n$. Then for linearly independent $v_1,\dots,v_m\in V$, we have

$$\frac{\omega(v_1,\dots,v_m)}{[\![\langle v_1,\dots,v_m\rangle]\!]}=\sum_{\alpha\in\mathcal{A}(n,m)}\omega_\alpha\frac{\det\langle v_1,\dots,v_m\rangle_\alpha}{[\![\langle v_1,\dots,v_m\rangle]\!]}$$

and

$$\frac{|\omega(v_1,\dots,v_m)|}{[\![\langle v_1,\dots,v_m\rangle]\!]}\leq\|\omega\|_{\Lambda^m}.$$

Definition. Let $M\subset\mathbb{R}^n$ be a C^1 manifold of dimension $1\leq m\leq n$ with orientation $\mathcal{O}$.

1. The **volume form** on M is the map $\tau_M:M\to\Lambda^m(\mathbb{R}^n)$ given by

$$\tau_M(z)=\frac{1}{[\![\langle v_1,\dots,v_m\rangle]\!]}\sum_{\alpha\in\mathcal{A}(m,n)}\det\langle v_1,\dots,v_m\rangle_\alpha e^\alpha,$$

where $(v_1,\dots,v_m)$ is any ordered basis in the orientation class induced on T_zM.

2. Suppose $\omega : M \to \Lambda^m(\mathbb{R}^n)$. We define the **density function** associated to ω on M to be $f_\omega : W \to \mathbb{R}$ by
$$f_\omega(z) = \langle \tau_M(z), \omega(z)\rangle_{\Lambda^m} = \frac{\omega(v_1, \dots, v_m)}{[\![\langle v_1, \dots, v_m\rangle]\!]},$$
where $(v_1, \dots, v_m)$ is any ordered basis in the orientation class induced on $T_z M$.

Proposition. Let $M \subset \mathbb{R}^n$ be a C^k m-manifold with orientation $\mathcal{O}$. The following holds:

1. Let $\tau_M : M \to \Lambda^m(\mathbb{R}^n)$ be the volume form on M. Then, $\|\tau_M(z)\|_{\Lambda^m} = 1$ for all $z \in M$ and if $(U, V, \phi) \in \mathcal{O}$, then
$$\tau_M \circ \phi = \frac{1}{[\![D\phi]\!]} \sum_{\alpha \in \mathcal{A}(n,m)} \det(D\phi)_\alpha e^\alpha.$$
Moreover,
$$\phi^* \tau_M = [\![D\phi]\!]\, e^1 \wedge \cdots \wedge e^m = J_\phi e^1 \wedge \cdots \wedge e^m.$$
In particular, τ_M is C^{k-1}.
2. If $\omega : M \to \Lambda^m(\mathbb{R}^n)$ is C^r for $0 \le r \le k-1$, then the density function $f_\omega : M \to \mathbb{R}$ is also C^r and hence $\mathcal{H}^m$-measurable on M.

Definition. Let $M \subset \mathbb{R}^n$ be an orientated C^1 manifold of dimension $1 \le m \le n$. We say that ω is M**-integrable** if $f_\omega : M \to \mathbb{R}$ is $\mathcal{H}^m$-integrable on M, in which case we define
$$\int_M \omega = \int_M f_\omega \, d\mathcal{H}^m.$$

Proposition. Let $M \subset \mathbb{R}^n$ be a C^k m-manifold with orientation $\mathcal{O}$. Let $\omega \in C^0(\underline{M; \Lambda^m}(\mathbb{R}^n))$ and suppose $\mathrm{supp}(\omega) \cap M \subset U \cap M$ for some $(U, V, \phi) \in \mathcal{O}$, where $\mathrm{supp}(\omega) = \overline{\mathrm{spt}(\omega)}$ is the support of ω. Then ω is M-integrable if and only if $\phi^*\omega$ is V-integrable, and
$$\int_M \omega = \int_V \phi^* \omega = \int_V \omega \circ \phi(\partial_1 \phi, \dots, \partial_m \phi)\, d\lambda.$$

Proof. We have
$$\int_M \omega = \int_M f_\omega \, d\mathcal{H}^m = \int_{M \cap U} f_\omega \, d\mathcal{H}^m = \int_{\phi(V)} f_\omega \, d\mathcal{H}^m.$$
However, by change of variable, we have
$$\int_{\phi(V)} f_\omega \, d\mathcal{H}^m = \int_V f_\omega \circ \phi J_\phi \, d\mathcal{H}^m.$$
Meanwhile,
$$f_\omega \circ \phi = \frac{\omega \circ \phi(\partial_1 \phi, \dots, \partial_m \phi)}{[\![D\phi]\!]}$$

and $J_\phi = [\![D\phi]\!]$. Therefore,

$$\int_M \omega = \int_V \phi^* \omega = \int_V \omega \circ \phi(\partial_1 \phi, \dots, \partial_m \phi)\, d\lambda.$$

□

4.3.2 Parition of unity

Definition. Suppose U is an open set in a normed vector space V, we write

$$C_c^\infty(U; V) = \{f \in C^\infty(U; V) : \operatorname{supp}(f) \text{ is compact}\},$$

where

$$\operatorname{supp}(f) = \overline{\{x \in U : f(x) \neq 0\}}.$$

Theorem (smooth partition of unity)**.** Let $\emptyset \neq U \subset \mathbb{R}^n$ be open and $A \neq \emptyset$ a set. Suppose $\{U_\alpha\}_{\alpha \in A}$ is an open cover of U. Then there exists $\{\psi_k\}_{k=0}^\infty \subset C_c^\infty(\mathbb{R}^n)$ such that the following holds:

1. $0 \leq \psi_k \leq 1$ for all $k \in \mathbb{N}$.
2. For each $k \in \mathbb{N}$ there exists $\alpha_k \in A$ such that $\operatorname{supp}(\psi_k) \subset U_{\alpha_k}$.
3. The sequence $\{\psi_k\}_{k=0}^\infty$ is locally finite: for each compact set $K \subset U$, the set $\{k \in \mathbb{N} : \operatorname{supp}(\psi_k) \cap K \neq \emptyset\}$ is finite.
4. We have that $\sum_{k=0}^\infty \psi_k = 1$ on U, where the convergence is uniform on compact sets, as the sum is finite in a neighborhood of each point in U.
5. If $K \subset U$ is compact, then there exists $N \in \mathbb{N}$ and an open set $K \subset V \subset U$ such that $\sum_{k=0}^N \psi_k = 1$ on V.

4.3.3 Stokes-Cartan theorem

Theorem (Stokes-Cartan theorem with boundary)**.** Suppose that $M \subset \mathbb{R}^n$ is an oriented C^2 manifold of dimension $1 \leq m \leq n$ with boundary. Let $\omega \in C^1(M; \Lambda^{m-1}(\mathbb{R}^n))$, and suppose that one of the following holds:

1. $\operatorname{supp}(\omega) \cap M$ is compact.
2. M is a closed subset of $\mathbb{R}^n$, ω is ∂M-integrable, $d\omega$ is M°-integrable, and $\|\omega\|_{\Lambda^{m-1}}$ is $\mathcal{H}^m$-integrable on M°.

Then,

$$\int_{M^\circ} d\omega = \int_{\partial M} \omega.$$

Remark. Our proof will use crucially the fact that M is a C^2 manifold. In fact the theorem is true for M being C^1. However, the proof needs more analytic tools.

Proof. We proceed the proof for this theorem with several steps.

Step 1: Computations in half-spaces

Suppose first $m \geq 2$ and let $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x_m > 0\}$, which is open and has boundary $\partial\mathbb{R}^m_+ = \mathbb{R}^{m-1} \times \{0\}$. Endow $\mathbb{R}^m_+$ with the natural orientation and the corresponding volume form

$$\tau_{\mathbb{R}^m_+} = e^1 \wedge \cdots \wedge e^m.$$

Also endow $\partial\mathbb{R}^m_+$ with the volume form

$$\tau_{\partial\mathbb{R}^m_+} = e^1 \wedge \cdots \wedge e^{m-1},$$

which corresponds to the natural orientation of $\mathbb{R}^{m-1}$.

Suppose $\omega \in C^1(\overline{\mathbb{R}^m_+}; \Lambda^{m-1}(\mathbb{R}^m))$ is such that $\operatorname{supp}(\omega)$ is compact. Then there exists $k > 0$ such that $\operatorname{supp}(\omega) \subset R_k = [-k, k]^{m-1} \times [0, k]$. Write

$$\omega = \sum_{i=1}^m \omega_i \widehat{e^i},$$

where $\widehat{e^i}$ denotes the $m-1$ form with e^i removed. It follows that

$$d\omega = \left[\sum_{i=1}^m (-1)^{i-1} \partial_i \omega_i\right] e^1 \wedge \cdots \wedge e^m.$$

Then we have

$$\int_{\mathbb{R}^m_+} d\omega = \int_{\mathbb{R}^m_+} \sum_{i=1}^m (-1)^{i-1} \partial_i \omega_i = \sum_{i=1}^m (-1)^{i-1} \int_{R_k} \partial_i \omega_i.$$

For $1 \leq i \leq m-1$, we have

$$\begin{aligned}
&\int_{-k}^k \partial_i \omega_i(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_m)\, dt \\
&= \omega_i(x_1, \ldots, x_{i-1}, k, x_{i+1}, \ldots, x_m) - \omega_i(x_1, \ldots, x_{i-1}, -k, x_{i+1}, \ldots, x_m) \\
&= 0
\end{aligned}$$

by the fundamental theorem of calculus. Fubini-Tonelli then implies that $\int_{R_k} \partial_i \omega_i = 0$ for all $1 \leq i \leq m-1$. On the other hand, for $i = m$, we similarly have

$$\begin{aligned}
&\int_0^k \partial_m \omega_m(x_1, \ldots, x_{m-1}, t)\, dt \\
&= \omega_m(x_1, \ldots, x_{m-1}, k) - (x_1, \ldots, x_{m-1}, 0) \\
&= -(x_1, \ldots, x_{m-1}, 0).
\end{aligned}$$

It follows that

$$\int_{R_k} \partial_m \omega_m = -\int_{[-k,k]^{m-1}} \omega_m(\cdot, 0) = -\int_{\mathbb{R}^{m-1}} \omega_m(\cdot, 0).$$

This implies that

$$\int_{\mathbb{R}^m_+} d\omega = (-1)^m \int_{\mathbb{R}^{m-1}} \omega_m(\cdot, 0)\, d\lambda_{m-1}.$$

However,

$$\int_{\mathbb{R}^{m-1}} \omega_m(\cdot, 0)\, d\lambda_{m-1} = \int_{\partial\mathbb{R}^m_+} \langle \tau_{\partial\mathbb{R}^m_+}, \omega\rangle_{\Lambda^{m-1}}\, d\mathcal{H}^{m-1} = \int_{\partial\mathbb{R}^m_+} \omega.$$

Therefore, for this choice of orientation for $\mathbb{R}^m_+$ and $\partial\mathbb{R}^m_+$, we have

$$\int_{\mathbb{R}^m_+} d\omega = (-1)^m \int_{\partial\mathbb{R}^m_+} \omega.$$

The case for $m = 1$ is similar to the previous argument with minor modifications.

Step 2: Computations in open sets

Suppose $\emptyset \neq V \subset \mathbb{R}^m$ is open and $\omega \in C^1(V; \Lambda^{m-1}(\mathbb{R}^m))$ is such that $\operatorname{supp}(\omega) \subset V$ is compact. It follows from a similar arguement as in Step 1 that

$$\int_V d\omega = 0.$$

Step 3: General M and special ω

Suppose $\operatorname{supp}(\omega) \subset U \cap M$, where $(U, V, \phi) \in \mathcal{O}$, the orientation of M.

If (U, V, ϕ) is local boundary coordinates, then

$$\int_{M^\circ} d\omega = \int_{\mathbb{R}^m_+} \phi^*(d\omega) = \int_{\mathbb{R}^m_+} d(\phi^*\omega) = (-1)^m \int_{\mathbb{R}^{m-1}} \phi^*\omega,$$

where $\phi^*(d\omega) = d(\phi^*\omega)$ relies on the fact that ϕ is C^2. Now recall that with an orientation on M°, the induced orientation is positive when m is even and negative when m is odd. It follows that

$$\int_{M^\circ} d\omega = \int_{\partial M} \omega.$$

On the other hand, if (U, V, ϕ) is not boundary coordinates, then Step 2 tells us that

$$\int_{M^\circ} d\omega = \int_{\mathbb{R}^m} d(\phi^*\omega) = 0.$$

Step 4: General M and ω from assumption 1

Suppose $\operatorname{supp}(\omega) \cap M$ is compact and denote the orientation $\mathcal{O} = \{(U_\alpha, V_\alpha, \phi_\alpha)\}_{\alpha\in A}$. Note that $\{U_\alpha\}_{\alpha\in A}$ is an open cover of M. Pick smooth partition of unity $\{\psi_k\}_{k=0}^\infty$ subordinate to $\{U_\alpha\}_{\alpha\in A}$.

Since $\operatorname{supp}(\omega) \cap M$ is compact, there exists $K \in \mathbb{N}$ such that $\sum_{k=0}^{K} \psi_k = 1$ on $\operatorname{supp}(\omega) \cap M$. Therefore, from Step 3 we have

$$\int_{\partial M} \omega = \int_{\partial M} \sum_{k=0}^{K} \psi_k \omega = \sum_{k=0}^{K} \int_{\partial M} \psi_k \omega = \sum_{k=0}^{K} \int_{M^\circ} d(\psi_k \omega) = \int_{M^\circ} d\left[\sum_{k}^{K} \psi_k \omega\right] = \int_{M^\circ} d\omega.$$

Step 5: General M and ω from assumption 2

Suppose ω satisfies the assumptions in item 2. Let $\phi \in C^1(\mathbb{R}^n;\mathbb{R})$ be such that $\phi = 1$ in $B(0,1)$ and $\phi = 0$ in $B(0,2)^c$. For $r > 0$ let $\phi_r \in C^1(\mathbb{R}^n;\mathbb{R})$ be given by $\phi_r(x) = \phi(\frac{x}{r})$. Since M is closed subset of $\mathbb{R}^n$, $B[0,2r] \cap M$ is compact for every $r > 0$, and thus $\phi_r\omega$ satisfies the assumptions in item 1. We then know that

$$\int_{\partial M} \phi_r \omega = \int_{M^\circ} d(\phi_r \omega) = \int_{M^\circ} \phi_r \, d\omega + \int_{M^\circ} d\phi_r \wedge \omega.$$

However, we have

$$\int_{\partial M} \phi_r \omega = \int_{\partial M} \phi_r f_\omega \, d\mathcal{H}^{m-1} \quad \text{and} \quad \int_{M^\circ} \phi_r d\omega = \int_{M^\circ} \phi_r f_{d\omega} \, d\mathcal{H}^{m}.$$

By assumption, we can use dominated convergence theorem to obtain

$$\int_{\partial M} \omega = \lim_{r\to\infty} \int_{\partial M} \phi_r \omega \quad \text{and} \quad \int_{M^\circ} d\omega = \lim_{r\to\infty} \int_{M^\circ} \phi_r d\omega.$$

On the other hand, it is easy to show that there exists constant $C > 0$ such that

$$\|d\phi_r \wedge \omega\|_{\Lambda^m} \leq \frac{C}{r} \|\omega\|_{\Lambda^{m-1}}$$

for all $r > 0$. Therefore,

$$\lim_{r\to\infty} \int_{M^\circ} d\phi_r \wedge \omega = 0.$$

It follows that

$$\int_{\partial M} \omega = \int_{M^\circ} \omega,$$

and the proof is complete. □

Theorem (Stokes-Cartan theorem without boundary)**.** Suppose that $M \subset \mathbb{R}^n$ is an oriented C^2 manifold. Let $\omega \in C^1(M;\Lambda^{m-1}(\mathbb{R}^n))$, and suppose that one of the following holds:

1. $\operatorname{supp}(\omega) \cap M$ is compact.
2. M is a closed subset of $\mathbb{R}^n$, $d\omega$ is M-integrable, and $\|\omega\|_{\Lambda^{m-1}}$ is $\mathcal{H}^m$-integrable on M.

Then,

$$\int_M d\omega = 0.$$

4.3.4 Fundamental theorems of vector analysis

We begin with integrals along curves in $\mathbb{R}^n$. Our first result establishes a connection between the integral of the tangential component of a vector field and the integral of an associated one-form.

Proposition. Let $M \subset \mathbb{R}^n$ be a C^k 1-manifold equipped with orientation $\mathcal{O} = \{(U_\alpha, V_\alpha, \phi_\alpha)\}_\alpha$. Let τ be the associated unit tangent vector field given by

$$\tau(z) = \frac{\phi'_\alpha(\phi_\alpha^{-1}(z))}{\left|\phi'_\alpha(\phi_\alpha^{-1}(z))\right|}.$$

Suppose $v \in C^0(M; \mathbb{R}^n)$ and $\omega \in C^0(M; \Lambda^1(\mathbb{R}^n))$ are related via

$$\omega(z) = \sum_{i=1}^{n} v_i(z) e^i.$$

Further suppose that ω is M-integrable. Then $v \cdot \tau$ is $\mathcal{H}^1$-integrable on M, and we have that

$$\int_M \omega = \int_M v \cdot \tau \, d\mathcal{H}^1.$$

Theorem (Fundamental theorem of line integrals)**.** Let $M \subset \mathbb{R}^n$ be an orientated C^2 1-manifold with boundary. Let τ denote the unit tangent vector field on M induced by orientation. Suppose also that $f \in C^1(W; \mathbb{R})$ where $W \subset \mathbb{R}^n$ is an open set such that $M \subset W$. Further suppose one of the following holds:

1. $\operatorname{supp}(f) \cap M$ is compact.
2. M is a closed subset of $\mathbb{R}^n$, f is $\mathcal{H}^0$-integrable on ∂M, and both f and ∇f are $\mathcal{H}^1$-integrable on M°.

Then we have

$$\int_{M^\circ} \nabla f \cdot \tau \, d\mathcal{H}^1 = \sum_{z \in \partial M} \sigma(z) f(z),$$

where $\sigma : \partial M \to \{\pm 1\}$ is the induced orientation on ∂M.

Proof. By the previous proposition and Stokes-Cartan, we have

$$\int_{M^\circ} \nabla f \cdot \tau \, d\mathcal{H}^1 = \int_{M^\circ} df = \int_{\partial M} f = \sum_{z \in \partial M} \sigma(z) f(z).$$

□

Next we discuss the divergence theorem.

Proposition. Suppose that $M \subset \mathbb{R}^n$ is an orientated C^k $(n-1)$-manifold. Let ν be the C^{k-1} unit normal vector field on M. Suppose that $v \in C^0(M;\mathbb{R}^n)$ and $\omega \in C^0(M;\Lambda^{n-1}(\mathbb{R}^n))$ are related via

$$\omega(z) = \sum_{i=1}^{n}(-1)^{i-1}v_i(z)\widehat{e}^i.$$

Further suppose that ω is M-integrable. Then $v \cdot \nu$ is $\mathcal{H}^{n-1}$ integrable and

$$\int_M \omega = \int_M v \cdot \nu \, d\mathcal{H}^{n-1}.$$

Theorem (divergence theorem)**.** Suppose that $n \geq 2$ and $M \subset \mathbb{R}^n$ is a C^2 n-manifold with boundary. Let ν be the outward pointing unit normal vector field on ∂M. Suppose that $v \in C^1(U;\mathbb{R}^n)$ for an open set U such that $\overline{M} \subset U$ and one of the following holds:

1. $\operatorname{supp}(v) \cap M$ is compact.
2. M is a closed subset of $\mathbb{R}^n$, v is $\mathcal{H}^{n-1}$-integrable on ∂M and $\operatorname{div} v$ and v are $\mathcal{H}^n$-integrable on M°.

Then,

$$\int_{M^\circ} \operatorname{div} v \, d\lambda = \int_{\partial M} v \cdot \nu \, d\mathcal{H}^{n-1}.$$

Proof. Equip M with the natural orientation and note that a previous proposition tells us that $\nu = n^{\text{out}}$. Define $\omega \in C^1(M;\Lambda^{n-1}(\mathbb{R}^n))$ by

$$\omega(z) = \sum_{i=1}^{n}(-1)^{i-1}v_i(z)\widehat{e}^i.$$

Then we know

$$d\omega = (\operatorname{div} v)e^1 \wedge \cdots \wedge e^n$$

and thus

$$\int_{M^\circ} d\omega = \int_{M^\circ} \operatorname{div} v \, d\lambda.$$

Now the previous proposition and the previous proposition implies that

$$\int_{M^\circ} d\omega = \int_{\partial M} \omega = \int_{\partial M} v \cdot \nu \, d\mathcal{H}^{n-1}.$$

□

We conclude our discussion of the fundamental theorems of vector analysis with Stokes' theorem.

Theorem (Stokes' theorem in 3D)**.** Suppose that $M \subset \mathbb{R}^3$ is an orientated C^2 2-manifold with boundary. Let ν be the unit normal C^1 vector field on M, and let τ be the C^1 unit tangent vector field on ∂M. Suppose that $v \in C^1(U;\mathbb{R}^3)$ for an open set U such that $\overline{M} \subset U$ and one of the following holds:

1. $\operatorname{supp}(v) \cap M$ is compact.
2. M is a closed subset of $\mathbb{R}^n$, v is $\mathcal{H}^1$-integrable on ∂M and v and $\operatorname{curl} v$ are $\mathcal{H}^2$-integrable on M°.

Then,

$$\int_{M^\circ} \operatorname{curl} v \cdot \nu = \int_{\partial M} v \cdot \tau.$$

Proof. Let $\omega \in C^1(M; \Lambda^1(\mathbb{R}^n))$ be defined by $\omega(z) = \sum_{i=1}^3 v_i e^i$. Then,

$$d\omega = (\operatorname{curl} v)_1 \, e^2 \wedge e^3 - (\operatorname{curl} v)_2 \, e^1 \wedge e^3 + (\operatorname{curl} v)_3 \, e^1 \wedge e^2.$$

Then previous proposition tells us

$$\int_{M^\circ} d\omega = \int_{M^\circ} (\operatorname{curl} v) \cdot \nu \, d\mathcal{H}^2.$$

On the other hand, Stokes-Cartan and previous proposition implies that

$$\int_{M^\circ} d\omega = \int_{\partial M} \omega = \int_{\partial M} v \cdot \tau d\mathcal{H}^1.$$

□