Mathematical Studies Analysis

Notes taken by Runqiu Ye
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 $0 \le i \le n$, 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 $0 \le i \le n$, $B[x_{i+1}, \varepsilon_{i+1}] \subset B[x_i, \varepsilon_i]$ for all $0 \le i < n$. 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) $\implies$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (5) $\implies$ (1), keeping in mind the following suggestions.

1. For (1) $\implies$ (2): Study the sets $C_n = \overline{T(B_X(0,n))} \subset Y$ for $n \geq 1$.
2. For (2) $\implies$ (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) $\implies$ (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 $0 \leq j \leq m$ 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

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

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 $\|x\| \leq 2$. 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

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

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

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

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. $T : X \to$ range(T) is a linear homeomorphism.
3. There exists $C \ge 0$ such that $\|x\|_X \le C \|Tx\|_Y$ 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, $T : X \to$ 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 contant $C \ge 0$ such that for each $y \in$ range(T) there exists a unique $x \in X$ such that $Tx = y$ and $\|x\| \le C \|y\| = C \|Tx\|$. Since T is a bijection, $\|x\| \le C \|Tx\|$ for all $x \in X$.

(3) $\implies$ (1). Let $x \in X$ be such that $Tx = 0$. It follows that $\|x\| \le C \|Tx\| = 0$. Therefore, $x = 0$ and T is injective. To show that range(T) is closed, consider a convergent sequence $\{y_n\} \subset$ range(T) with $y_n = Tx_n$. Since for any $n, m \in \mathbb{N}$ we have

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

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

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

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 $\|x_0\| \le 2\alpha \|y_0\|$. Suppose we have selected y_i, x_i for $0 \le i \le n$. Set $y_{n+1} = y_n - Sx_n$ and select x_{n+1} be such that $Tx_{n+1} = y_{n+1}$ and $\|x_{n+1}\| \le 2\alpha \|y_{n+1}\|$. Then, we have

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

   and

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

   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 $\|x\| \le C\|Tx\|$ for all $x \in X$. Now supose $S \in \mathcal{L}(X;Y)$ is such that $\|T - S\| < (2C)^{-1}$. Claim that S is also injective with closed range. Indeed,

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

   This shows that $\|x\| \le 2C\|Sx\|$ 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. This directly follows from

   $$\mathcal{H}(X;Y) = \{T \in \mathcal{L}(X;Y) : T \text{ is surjective}\} \cap \{T \in \mathcal{L}(X;Y) : T \text{ is injective with closed range}\}.$$

□

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

$\square$

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 $-p(-x_0) \leq$ LHS and RHS $\leq p(x_0)$, so LHS, RHS $< \infty$. Now pick $v \in \mathbb{R}$ such that LHS $\leq v \leq$ 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 exsits 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 theoerm.

□

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. $\text{range}(J) \subset X^{**}$ is a norming set for X^*.
4. X is Banach if and only if $\text{range}(J)$ is closed.

Proof. Note that we have

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

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 $\text{range}(J) \subset X^{**}$. Therefore, X is Banach if and only if $\text{range}(J)$ is Banach. However, $X^{**} = \mathcal{L}(X^*, \mathbb{F})$ is Banach, so $\text{range}(J)$ is Banach if and only if $\text{range}(J)$ is closed. This implies (4).

To show (3), note that we have

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

This shows (3), completing the proof. □

2 Differential Calculus

2.1 Inverse and implicit function 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 $0 \le r \le r_0$.

Proof. ***** TO-DO *****

□

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 (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 ***** □

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 ***** □

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** $\bigoplus_\alpha \mathfrak{M}_\alpha$ 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. $\bigoplus_\alpha \mathfrak{M}_\alpha \subset \sigma(\mathcal{R})$. If A countable then $\sigma(\mathcal{R}) = \bigoplus_\alpha \mathfrak{M}_\alpha$.

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 $\bigoplus_\alpha \mathfrak{M}_\alpha = \sigma(\mathcal{E})$. Moreover, if A is countable and $X_\alpha \in \mathcal{E}_\alpha$ for all $\alpha \in A$, then $\bigoplus_\alpha \mathfrak{M}_\alpha$ 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 $\bigoplus_\alpha \mathfrak{M}_\alpha \subset \sigma(\mathcal{R})$.

On the other hand, if A is countable, then

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

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

2. It is clear that $\sigma(\mathcal{E}) \subset \bigoplus_\alpha \mathfrak{M}_\alpha$. 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 $\bigoplus_\alpha \mathfrak{M}_\alpha \subset \sigma(\mathcal{E})$.

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.

□

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

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

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

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

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 $\bigoplus_{i=1}^n \mathfrak{B}_{X_i}$ is generated by the same set. Therefore, $\bigoplus_{i=1}^n \mathfrak{B}_{X_i} = \mathfrak{B}_X$.

□

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 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 $\sum_{n=0}^\infty \gamma(E_{m,n}) \le \mu^*_{2^{-m}}(A) + 2^{-m}$. 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

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

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

$$\mu^*_d(E) \le \mu^*_d(A) \le \mu^*_d(E),$$

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 $\mu^*_d(A_n) \ge \mu^*_\delta(A_n)$ 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

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

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

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

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

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

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

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

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

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

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

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

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 $\{\varphi_k\}_{k=0}^{\infty} \subset S(X; \overline{\mathbb{R}})$ such that $\varphi_k \to f$ 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 $0 \leq \varphi_k \leq \varphi_{k+1} \leq f$.
- On the set $\{f < 0\}$, we have $f \leq \varphi_{k+1} \leq \varphi_k \leq 0$.
- If f is actually from X to $\mathbb{R}$ and is bounded, then $\varphi_k \to f$ 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 $\varphi_k : [0, \infty)$ by

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

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

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

In the general case, apply the special case to f on $\{f \ge 0\}$ 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 $\{\varphi_k\}_{k=0}^{\infty} \in S(X;Y)$ such that $\varphi_k \to f$ pointwise.

Proof. (2) $\implies$ (1). The pointwise limit of measurable function is measurable. On the other hand, $f(X) = \overline{\bigcup_{k=0}^{\infty} \varphi_k(X)}$, which is separable since $\varphi_k(X)$ 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 $y_0^n, \dots, y_{K(n)}^n \in Y$ 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 $k \ge 1$ 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 $M(n) < k \le K(n)$.

Define $\varphi_n : Y \to \{y_0^n, \dots, y_{M(n)}^n\}$ via $\varphi_n(y) = y_k^n$ if $y \in V_k^n$. Clearly φ_n is simple and $d(\varphi_n(y), y) < 2^{-n}$ for all $n \in \mathbb{N}$ and $y \in Y$. Therefore, $\varphi_n(y) \to (y)$ pointwise. Then $f_n = \varphi_n \circ f$ are simple functions from X to Y. Also, since $\varphi_n \to \mathrm{id}$ 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 $\{\varphi_n\}_{n=0}^{\infty} \subset S(X;Z)$ such that $\varphi_n \to h \circ f$ pointwise. Then, $h^{-1} \circ \varphi_n \in S(X;Y)$ is such that $h^{-1} \circ \varphi_n \to h^{-1} \circ h \circ f = f$ 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 $\{\varphi_n\}_{n=0}^{\infty} \in S(N^c; Y)$ such that $\varphi_n \to F = f$ 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 $\mu(\operatorname{supp}(\psi)) < \infty$ where $\operatorname{supp}(\psi) = \{x \in X : \psi(x) \neq 0\}$ is the 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 theoerm.

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 $\{\varphi_n\}_{n=0}^{\infty} \subset S(X; Y)$ be such that $\varphi_n \to g$ 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 $f = \lim_{n\to\infty} \varphi_n$. This implies that $g = \lim_{n\to\infty} \varphi_n$ on $(N \cup Z)^c$.
2. Same proof as the first item but let $\{\varphi_n\}_{n=0}^{\infty} \in S_{\text{fin}}(X; Y)$.

□

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 $\varphi_n = \chi_{X_n} \psi_n$. 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 $\{\varphi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; Y)$ such that $\varphi_n \to f$ pointwise on N^c where N is null.

Pick $\varepsilon > 0$ such that $\{0\} \cap B(y, \varepsilon) = \emptyset$. Set $X_n = \varphi_n^{-1}(B(y, \varepsilon))$. Then we have $\mu(X_n) < \infty$. For any $x \in N^c$ and n sufficiently large, $\varphi_n(x) \in B(y, \varepsilon)$. 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 $\{\varphi_n\}_{n=0}^\infty \subset S(X;V)$ such that $\|\varphi_n\| \leq 2\|f\|$ on X such that $\varphi_n \to f$ pointwise a.e. as $n \to \infty$. Moreover, the same equivalence holds with μ-measurablility replaced by strongly μ-measurablility and $\{\varphi_n\}_{n=0}^\infty$ replaced by $\{\varphi_n\}_{n=0}^\infty \subset S_{\text{fin}}(X;V)$.

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 $\|S_n(u)\| \le 2\|u\|$. 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 $0 \le k \le n$, 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 $\varphi_n \in S(X;V)$ by

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

Then, $\|\varphi_n\| \le 2\|f\|$ and $\varphi_n(x) = \psi_n(x) \to f(x)$ as $n \to \infty$ for $x \in N^c$. Therefore, $\varphi_n \to f$ 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
$$\left\|\int_X f \, d\mu\right\| \le \int_X \|f\| \, d\mu.$$
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 $\overline{\mathbb{R}}$-valued functions

Note that if $(X, \mathfrak{M}, \mu)$ is a measure space and $\varphi \in S(X; [0, \infty])$, then
$$\int_X \varphi\, d\mu = \sup\left\{\int_X \psi\, d\mu : \psi \in S(X; [0, \infty]) \text{ and } \psi \le \varphi \text{ a.e.}\right\}.$$

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 $\int_X f \, d\mu$. □

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 $\varphi \in S(X; [0, \infty])$ such that $\varphi \leq f$ a.e. and $\alpha \in (0, 1)$. Let $N \in \mathfrak{M}$ be a null set and $\varphi \leq f$ on N^c. Also, for each $n \in \mathbb{N}$, let $E_n = \{x \in X : f_n(x) \geq \alpha\varphi(x)\}$. 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
$$\alpha \int_{N\cup E_n} \varphi \, d\mu = \int_{E_n} \alpha\varphi \, d\mu \leq \int_{E_n} f_n \, d\mu \leq \int_X f_n \, d\mu$$
4. We have
$$\int_X \varphi \, d\mu = \lim_{n\to\infty} \int_{N\cap E_n} \varphi \, d\mu.$$

Therefore,
$$\alpha \int_X \varphi \, d\mu = \lim_{n\to\infty} \alpha \int_{N\cup E_n} \varphi \, d\mu \leq \lim_{n\to\infty} \int_X f_n \, d\mu.$$
Since the above inequality holds for all $\alpha \in (0, 1)$, we konw $\int_X \varphi \, d\mu \leq \lim_{n\to\infty} \int_X f_n \, d\mu$. This is then true for all simple function φ such that $\varphi \leq f$ a.e. Taking the sup gives
$$\int_X f \, d\mu \leq \lim_{n\to\infty} f_n \, d\mu.$$
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 $\{\varphi_n\}_{n=0}^{\infty}, \{\psi_n\}_{n=0}^{\infty} \subset S(X; [0, \infty])$ such that $\lim_{n\to\infty} \varphi_n = F$ and $\lim_{n\to\infty} \psi_n = G$, $0 \le \varphi_n \le \varphi_{n+1} \le F$ and $0 \le \psi_n \le \psi_{n+1} \le G$. Then

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

It follows then from monotone convergence theorem that

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

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 $0 \le f^{\pm} \le |f|$, 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 $f \le g$ a.e. then $\int_X f\ d\mu \le \int_X g\ d\mu$.

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 inequality)**.** If f is measurable, 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 $\text{supp}(f) \subset E$.
2. If f is extended integrable, then there exsits 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 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, $\text{supp}(f) = \bigcup_{k=0}^{\infty} \left\{|f| > 2^{-k}\right\}$, but

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

It follows that $\text{supp}(f)$ 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 E 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 $g \leq f_k$ 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 $g \geq f_k$ 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 $\operatorname{supp}(f_k^-)$ 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, $g \leq f_k$ a.e. implies $g \leq f$ a.e. as well. It follows that $-f_k \leq -g$ and $-f \leq -g$. Therefore, $f_k^- \leq g^-$ and $f^- \leq g^-$. This shows that f_k and f are extended-integrable. Next, note that

$$\int_X g \, d\mu \leq \int_X \inf_{j\geq k} f_j \, d\mu \leq \int_X f_k \, d\mu.$$

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

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

□

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 $|f_k| \leq g_k$ 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 $\varphi : X \to V$ simple. Note then $\|\varphi\| : X \to [0, \infty)$ is a simple function now. Then, φ is a **finite** simple function if and only if $\|\varphi\|$ is integrable.

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

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

$$\mu(\mathrm{supp}(\varphi)) = \mu(\{x \in X : \|\varphi(x)\| > 0\}) = \mu(\{\|\varphi\| \ge m\}).$$

By Chebyshev inequality, we have

$$\mu(\mathrm{supp}(\varphi)) \le \frac{1}{m} \int_X \|\varphi\| \ d\mu < \infty.$$

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 $\left\{\varphi_k^j\right\}_{k=0}^{\infty} \subset S_{\text{fin}}(X; V)$ such that

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

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

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

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

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

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

On the other hand,

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

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 $\{\varphi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; V)$ such that $\varphi_n \to f$ a.e. and

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

in which case we define

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

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 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. ($\Longrightarrow$) 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 $\{\varphi_n\}_{n=0}^{\infty} \subset S_{\text{fin}}(X; V)$ is such that $\varphi_n \to f$ a.e. and $\int_X \|f - \varphi_n\| \to 0$. Then,

$$\int_X \|f\| \, d\mu \le \int_X \|f - \varphi_n\| \, d\mu + \int_X \|\varphi_n\| \, d\mu < \infty$$

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

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

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

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

$$\int_X \|f - \varphi_n\| \ d\mu \to 0$$

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

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

It follows then from the dominated convergence theorem that

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

□

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

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

Therefore, $T \circ f$ is integrable and

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

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 th 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 Constructing 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 $\nu \leq \mu$ 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 $\mu^* = \gamma$. 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

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

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 $\nu \le \mu$. Let $E \in \mathfrak{N} \subset \mathfrak{M}$ and $\{A_k\}_{k=0}^{\infty} \subset \mathfrak{A}$ that covers E. It follows that

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

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

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

Therefore, $\nu \le \mu$.

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,

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

It follows that $\mu\left(\bigcup_{k=0}^{\infty} A_k \setminus E\right) < \varepsilon$ and thus

$$\mu(E) \le \mu\left(\bigcup_{k=0}^{\infty} A_k\right) = \nu\left(\bigcup_{k=0}^{\infty} A_k\right) = \nu(E) + \nu\left(\bigcup_{k=0}^{\infty} A_k \setminus E\right) \le \nu(E)+\varepsilon,$$

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.

□

Theorem (Product measures)**.** Let $2 \le n \in \mathbb{N}$ and suppose $(X_i, \mathfrak{M}_i, \mu_i)$ is measure space for $1 \le i \le n$. Let $X = \prod_i X_i$ and

$$\mathcal{E} = \left\{E = \prod_i E_i : E_i \in \mathfrak{M}_i \text{ for } 1 \le i \le n\right\}.$$

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 $\left\{E^k\right\}_{k=0}^{\infty} \subset \mathcal{E}$ and $\left\{F^k\right\}_{k=0}^{\infty} \subset \mathcal{E}$ 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 $1 \le i \le n$, 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,
$$E^c = \prod_{i=1}^{n}\left(E_i^c \times \prod_{j\neq i} E_j\right) \in \mathcal{E}.$$
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 $x = (x_1, \dots, x_n) \in X$. Now fix $(x_2, \dots, x_n)$, 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 $\left\{F^j\right\}_{j=0}^J \subset \mathfrak{A}$ 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 $\left\{E^k\right\}_{k=0}^\infty \subset \mathcal{E}$ 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 $1 \le i \le n$, 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 $\left\{A^k\right\}_{k=0}^\infty$ 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.

$\square$

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 $\bigoplus_{i=1}^n \mathfrak{M}_i$. 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 $\nu : \bigoplus_{i=1}^n \mathfrak{M}_i \to [0, \infty]$ 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

We first need to develop a few facts in linear algebra.

Proposition. Let $V_1, \dots, V_n, W$ be vector space over $\mathbb{F}$ and $T \in L(V_1, \dots, V_n; W)$. Suppose $x_i^j \in V_i$ for $j = 0, 1$ and $1 \le i \le n$. Then,

$$\begin{aligned} T(x_1^0 + x_1^1, \dots, x_n^0 + x_n^1) &= \sum_{\beta \in B(n)} T(x^{\beta(1)}, \dots, x^{\beta(n)}) \\ &= \sum_{m=0}^{n} \sum_{\beta \in B_m(n)} T(x_1^{\beta(1)}, \dots, x_n^{\beta(n)}), \end{aligned}$$

where

$$\begin{aligned} B(n) &= \{\beta : \{1, \dots, n\} \to \{0, 1\}\}, \\ B_m(n) &= \left\{\beta \in B(n) : \sum \beta(k) = m\right\}. \end{aligned}$$

Proof. Induction on $n \ge 1$. □

Definition. 1. For $1 \le k \le n$ 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 $1 \le k \le n$, 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 $1 \le k \le n$ and $\beta \in B_k(n)$, we set $\alpha \in \mathcal{A}(n, k)$ to be an increasing enumeration of $\{1 \le i \le n : \beta(i) = 1\}$. 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 $\emptyset \neq U \subset \mathbb{R}^n$, $f \in C^1(U;\mathbb{R}^m)$ 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
$$\mathcal{H}^n(ME) = J_f(z)\lambda(E).$$
Then 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 $N_f(\cdot, E) : Y \to [0,\infty]$ by
$$N_f(y,E) = \mathcal{H}^0\left(E \cap f^{-1}(\{y\})\right) = \mathcal{H}^0\left(\{x \in E : f(x) = y\}\right).$$

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 $N_f(\cdot, E) : \mathbb{R}^m \to [0,\infty]$ is $\mathfrak{H}^n(\mathbb{R}^m)$ measurable.

Proof. Homework. □

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
$$\int_E J_f\, d\lambda = \int_{\mathbb{R}^m} N_f(\cdot, E)\, d\mathcal{H}^n.$$

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, $E = \bigcup_{k=0}^\infty B(x_k, r_k)$, and
$$(1-\varepsilon)^{n+1}\int_F J_f\, d\lambda \leq \mathcal{H}^n(f(E)) \leq (1+\varepsilon)^{n+1}\int_F J_f\, d\lambda$$
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
$$\mathcal{H}^n(f(E_k)) = \int_{\mathbb{R}^m} N_f(\cdot, E_k)\, d\mathcal{H}^n.$$
Summing the inequalities, we can then obtain from monotone convergence theoerm that
$$(1-\varepsilon)^{n+1}\int_E J_f\, d\lambda \leq \int_{\mathbb{R}^m}\sum_{k=0}^\infty N_f(\cdot, E_k)\, d\mathcal{H}^n = \int_{\mathbb{R}^m} N_f(\cdot, E)\, d\mathcal{H}^n \leq (1+\varepsilon)^{n+1}\int_E J_f\, d\lambda.$$

Since this holds for all $\varepsilon > 0$, we have

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} N_f(\cdot, E) \, d\mathcal{H}^n.$$

□

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 $\mathcal{H}^n(f(Z_k)) = 0$. □

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

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} N_f(\cdot, E) \, d\mathcal{H}^n = \int_{f(E)} N_f(\cdot, E) \, d\mathcal{H}^n.$$

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

$$\int_{V \cap E} J_f \, d\lambda = \int_{\mathbb{R}^m} N_f(\cdot, E \cap V) \, d\mathcal{H}^n.$$

On the other hand,

$$\int_{E \cap Z} J_f \, d\lambda = 0 = \int_{f(E \cap Z)} N_f(\cdot, E \cap Z) \, d\mathcal{H}^n = \int_{\mathbb{R}^m} N_f(\cdot, E \cap Z) \, d\mathcal{H}^n.$$

Adding the equality together gives

$$\int_E J_f \, d\lambda = \int_{\mathbb{R}^m} N_f(\cdot, E) \, d\mathcal{H}^n.$$

□