

Mathematical Studies Analysis

Notes taken by Runqiu Ye
Carnegie Mellon University

Spring 2025

Contents

1	Advanced topics in metric space theory	3
1.1	Baire category	3
1.2	Open mapping theorem	4
1.3	Hahn-Banach theorem and duality	8
2	Differential Calculus	12
2.1	Inverse and implicit function theorem	12
3	Measure and integration	14
3.1	Introduction to abstract measure theory	14
3.1.1	Outer measure and the Carathéodory construction	14
3.1.2	Constructing outer measures	14
3.2	Measurable and μ -measurable functions	17

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. \square

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 $i \in \mathbb{N}$, $B[x_i, \varepsilon_i] \subset U_i$, $B[x_{i+1}, \varepsilon_{i+1}] \subset B[x_i, \varepsilon_i] \subset B(x, \varepsilon)$, and $\varepsilon_i < 2^{-i}\varepsilon$.

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

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

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, $\bigcap_{n=1}^\infty C_n^c = (\bigcup_{n=1}^\infty C_n)^c = \emptyset$, 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 $x_1 \in B_X(0, 1)$ 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

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

It follows that

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

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 $\|x_0\| < 1$ such that $\|y - Tx_0\| < 2^{-1}R$. 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. Therefore, 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 $T : X_1 \rightarrow X_2$, where X_1 and X_2 are X equipped with norms $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively, be the identity map. Then for any $x \in X$ with $\|x\|_1 = 1$, we have

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

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

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

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

Theorem. Let X and Y be Banach spaces and let $T : X \rightarrow Y$ 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, $x_n \rightarrow x$ for some $x \in X$. Since $T \in \mathcal{L}(X; Y)$, we have

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

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

(b) \implies (a). Let $\pi_1 : \Gamma(T) \rightarrow X$ and $\pi_2 : \Gamma(T) \rightarrow Y$ 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. \square

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

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

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

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

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

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

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

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

and $y_n \rightarrow Tx$. Hence, $\text{range}(T)$ is closed and the proof is complete. \square

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

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

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

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

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

and

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

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

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

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

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

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

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

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

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

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

3. ***TO-DO***

□

1.3 Hahn-Banach theorem and duality

Theorem (Hahn-Banach theorem in \mathbb{R}). Let X be a real vector space and suppose $p : X \rightarrow \mathbb{R}$ 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 $l : Y \rightarrow \mathbb{R}$ is a linear map such that $l \leq p$ on Y . Then there exists linear map $L : X \rightarrow \mathbb{R}$ such that $L \leq p$ on X and $L = l$ on Y .

Proof. Let

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

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

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

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 that $M = X$.

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

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

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

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

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

Theorem (Hahn-Banach theorem in \mathbb{C}). Let X be complex vector space and suppose $p : X \rightarrow \mathbb{R}$ 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 $l : Y \rightarrow \mathbb{C}$ is a linear map such that $|l| \leq p$ on Y . Then there exists linear map $L : X \rightarrow \mathbb{C}$ such that $|L| \leq p$ on X and $L = l$ on Y .

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

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

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 $\Lambda : X \rightarrow \mathbb{R}$ that agrees with λ on Y .

Define $L : X \rightarrow \mathbb{C}$ 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

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

as desired. \square

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

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

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

1. There exists $\lambda \in X^*$ such that $\|\lambda\| = \|x\|$ and

$$\lambda(x) = \|\lambda\| \|x\| = \|x\|^2.$$

2. We have

$$\|x\| = \max_{\substack{w \in X^* \\ \|w\|=1}} |w(x)|.$$

3. $x = 0$ if and only if $w(x) = 0$ for all $w \in X^*$.

Proof. 1. Let $Y = \mathbb{F}x$ and define $\lambda \in Y^*$ by $\lambda(ax) = a \|x\|^2$. Apply Hahn-Banach theorem.

2. Suppose $x \neq 0$. Define $w = \frac{\lambda}{\|x\|}$ then it follows that $|w(x)| = \|x\|$.

3. Follows directly from (2). □

Proposition. Let X be normed vector space. Then the mapping $\langle \cdot, \cdot \rangle : X^* \times X \rightarrow \mathbb{F}$ 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 $J : X \rightarrow X^{**}$ 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 isomorphic 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 $f : U \rightarrow Y$ 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

$$(1 - \varepsilon) \|Df(z)(x - y)\| \leq \|f(x) - f(y)\| \leq (1 + \varepsilon) \|Df(z)(x - y)\|.$$

4. The restriction $f : B(z, r) \rightarrow f(B(z, r))$ is bi-Lipschitz homeomorphism.

Proof. Since $Df(z)$ injective with closed range, there exists $\theta > 0$ such that

$$\theta \|h\| \leq \|Df(z)h\|$$

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 $\|Df(z) - T\| < \delta$ 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

$$\|Df(x) - Df(z)\| < \min \{\delta, \theta\varepsilon\}.$$

In particular, $Df(x)$ is injective with closed range for all $x \in B[z, r]$. By the mean value theorem, for any $x, y \in B(x, r)$

$$\begin{aligned} \|f(x) - f(y) - Df(z)(x - y)\| &\leq \sup_{w \in B(z, r)} \|Df(w) - Df(z)\| \|x - y\| \\ &\leq \theta\varepsilon \|x - y\| \\ &\leq \varepsilon \|Df(z)(x - y)\|. \end{aligned}$$

It follows that

$$(1 - \varepsilon) \|Df(z)(x - y)\| \leq \|f(x) - f(y)\| \leq (1 + \varepsilon) \|Df(z)(x - y)\|,$$

as desired.

This also implies that

$$(1 - \varepsilon)\theta \|x - y\| \leq \|f(x) - f(y)\| \leq (1 + \varepsilon) \|Df(z)\| \|x - y\|,$$

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 $f : U \rightarrow Y$ differentiable with Df continuous at z . Suppose $Df(z) \in \mathcal{L}(X; Y)$ surjective. Then there exists $r_0, \gamma > 0$ such that

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

Proof. *** TO-DO *** □

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

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

$$\frac{1}{C_1} \leq \|Df^{-1}(y)\| \leq \frac{1}{C_0}$$

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 $f^{-1} \in C_b^k(f(V_k); X)$.

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 $f : U \rightarrow Z$ is differentiable in U with Df continuous at (x_0, y_0) . Further suppose $z_0 = f(x_0, y_0)$ and $D_2f(x_0, y_0) \in \mathcal{L}(Y; Z)$ 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. $D_2f(x, g(x, z))$ is an isomorphism for all $(x, z) \in V \times W$, and

$$\begin{aligned} D_1g(x, z) &= -[D_2f(x, g(x, z))]^{-1} D_1f(x, g(x, z)), \\ D_2g(x, z) &= [D_2f(x, g(x, z))]^{-1}. \end{aligned}$$

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 $g \in C_b^k$.

3 Measure and integration

3.1 Introduction to abstract measure theory

3.1.1 Outer measure and the Carathéodory construction

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}, μ) is a complete measure space.

Definition. 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 $\mu^*(A) = \mu(E)$.

3.1.2 Constructing outer measures

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

Theorem. Let X be a set and (\mathcal{E}, γ) be a gauge on X . Define $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ 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}, μ) , a complete measure space through Carathéodory construction.

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

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

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

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

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

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

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

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

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

as desired. □

Theorem. Let (X, d) be metric space with (\mathcal{E}, γ) 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}, γ) , or
2. $\mu^* = \mu_d^*$ is generated by $(\mathcal{E}_\delta, \gamma_\delta)$.

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

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

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

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

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

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

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

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

Step 3: further reduction.

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

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

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

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

Therefore, K is the desired F_σ set.

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

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

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

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

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

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

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

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

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

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

□

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

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

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

which are open for all $n, m \in \mathbb{N}$. Now notice that $G_n \subset Z_{n,m+1} \subset Z_{n,m} \subset V_n$. Note that $\mu^*(V_n) < \infty$, so $\mu^*(G_n) = \lim_{m \rightarrow \infty} \mu^*(Z_{n,m})$. 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 $\mu : \mathfrak{M} \rightarrow [0, \infty]$ is a measure. Say μ is outer-regular if

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

3.2 Measurable and μ -measurable functions

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

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

Proof. (2) \implies (1). The pointwise limit of measurable function is measurable. On the other hand, $f(X) = \bigcup_{k=0}^{\infty} \varphi_k(X)$, 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 $k \geq 1$ define $V_k^n = B(y_k^n, 2^{-n}) \setminus \bigcup_{j=0}^{k-1} B(y_j^n, 2^{-n})$. Then, $Y = \bigsqcup_{k=0}^{M(n)} V_k^n$ where $V_k^n = \emptyset$ for $M(n) < k \leq K(n)$.

Define $\varphi_n : Y \rightarrow \{y_0^n, \dots, y_{M(n)}^n\}$ 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, $\varphi_n(y) \rightarrow (y)$ pointwise. Then $f_n = \varphi_n \circ f$ are simple functions from X to Y . Also, since $\varphi_n \rightarrow \text{id}$ pointwise, $f_n \rightarrow f$ pointwise.

Now consider the general case in which $f(X)$ is a separable subset of Y . Then there exists a homeomorphism $h : f(X) \rightarrow Z$ 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 $h \circ f : X \rightarrow Z$ is measurable with Z totally bounded, so the special case provides a sequence $\{\varphi_n\}_{n=0}^{\infty} \subset S(X; Z)$ such that $\varphi_n \rightarrow h \circ f$ pointwise. Then, $h^{-1} \circ \varphi_n \in S(X; Y)$ is such that $h^{-1} \circ \varphi_n \rightarrow h^{-1} \circ h \circ f = f$ pointwise, using continuity of h and h^{-1} . □

Definition (Almost everywhere). Let (X, \mathfrak{M}, μ) 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 $x \in X \setminus N$.

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

1. There exists $\{\psi_n\}_{n=0}^{\infty} \subset S(X; Y)$ such that $\psi_n \rightarrow f$ a.e. in X .
2. There exists a measurable and separatedly valued $F : X \rightarrow Y$ such that $f = F$ a.e.

3. There exists a null set $N \in \mathfrak{M}$ and a measurable $F : X \rightarrow Y$ such that $f = F$ on $X \setminus N$ and $f(X \setminus N)$ is separable in Y .

Proof. (1) \implies (2). There exists $N \in \mathfrak{M}$ null such that $\psi_n \rightarrow f$ pointwise in $X \setminus N$. Thus, $f : X \setminus N \rightarrow Y$ 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 $F : X \rightarrow Y$ by

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

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

(2) \implies (3). Trivial.

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

□

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

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

1. We say $f : X \rightarrow Y$ is almost measurable if $f = F$ a.e. if $F : X \rightarrow Y$ is measurable.
2. We say $f : X \rightarrow Y$ is almost separably valued if there exists a null set $N \in \mathfrak{M}$ such that $f(X \setminus N)$ is separable.
3. We say $f : X \rightarrow Y$ 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 $f : X \rightarrow Y$ is strongly μ -measurable if there exists $\{\psi_n\}_{n=0}^\infty \subset S_{\text{fin}}(X; Y)$ such that $\psi_n \rightarrow f$ a.e. as $n \rightarrow \infty$.