INTRODUCTION TO ANALYSIS

CARSON JAMES

Contents

Preface

1. Real and Complex Numbers

Note 1.0.1. As a starting point, we will take as fact the existence of the **natural numbers**

$$\mathbb{N} = \{1, 2, \cdots\}$$

the **integers**

$$\mathbb{Z} = \{\cdots, -2, -2, 0, 1, 2, \cdots\}$$

and the **rational numbers**

$$\mathbb{Q} = \left\{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N}\right\}$$

1.1. Real Numbers.

Definition 1.1.1. Let X be a set and $\leq$ a relation on X. Then $\leq$ is said to be a **total order** if for each $a, b, c \in X$,

(1) $a \leq a$
(2) $a \leq b$ and $b \leq c$ implies that $a \leq c$
(3) $a \leq b$ and $b \leq a$ implies that $a = b$
(4) $a \leq b$ or $b \leq a$

Exercise 1.1.2. We define the relation $\leq$ on $\mathbb{Q}$ defined by

$$\frac{a}{b} \leq \frac{c}{d} \text{ iff } ad \leq bc$$

Then $\leq$ is a total order of $\mathbb{Q}$.

Proof. Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$. Then

(1) $\frac{a}{b} \leq \frac{a}{b}$ since $ab \leq ab$.
(2) if $\frac{a}{b} \leq \frac{c}{d}$ and $\frac{c}{d} \leq \frac{e}{f}$, then $ad \leq bc$ and $cf \leq de$. Multiplying the first inequality by f and the second inequality by b, we obtain $adf \leq bcf \leq bde$. Dividing both sides by d yields $af \leq be$. Hence $\frac{a}{b} \leq \frac{e}{f}$.
(3) if $\frac{a}{b} \leq \frac{c}{d}$ and $\frac{c}{d} \leq \frac{a}{b}$, then $ad \leq bc$ and $bc \leq ab$. This implies that $ad = bc$. Hence $\frac{a}{b} = \frac{c}{d}$.
(4)

$\square$

2. Metric Spaces

2.1. Introduction.

3. Topology

Definition 3.0.1. Let X be a topological space and $S, N \subset X$. Then N is said to be a **neighborhood** of S if there exists $U \subset X$ such that U is open and $S \subset U \subset N$. For $S \in X$, we denote the set of neighborhoods of S by $\mathcal{N}_S$

Exercise 3.0.2. Let X be a topological space and $A \subset X$. Then A is open iff for each $a \in A$, there exists $U_a \in \mathcal{N}_a$ such that U_a is an open of a and $U_a \subset A$.

Proof. Suppose that A is open. Let $a \in A$. Then $A \in \mathcal{N}_a$, A is an open and $A \subset A$. Conversely, suppose that or each $a \in A$, there exists $U_a \in \mathcal{N}_a$ such that U_a is open and $U_a \subset A$. Then $A = \bigcup_{a \in A} U_a$ is open. □

Definition 3.0.3. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be topological spaces and $f : X \to Y$. Then

1. f is said to be **continuous** if for each $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.
2. f is said to be open if for each $A \in \mathcal{A}$, $f(A) \in \mathcal{B}$.
3. f is said to be **closed** if for each $A \subset X$, if $A^c \in \mathcal{A}$, then $f(A)^c \in \mathcal{B}$.

Exercise 3.0.4. Let X, Y be topological spaces and $\phi : X \to Y$ a homeomorphism. Then for each $A \subset X$,

1. $\overline{\phi(A)} = \phi(\overline{A})$
2. $\phi(A)^\circ = \phi(A^\circ)$

Proof.

1. Let $A \subset X$. Since $A \subset \overline{A}$, we have that $\phi(A) \subset \phi(\overline{A})$. Since $\overline{A}$ is closed, $\phi(\overline{A})$ is closed and thus $\overline{\phi(A)} \subset \phi(\overline{A})$. Conversely, let $x \in \phi(\overline{A})$. Then $\phi^{-1}(x) \in \overline{A}$. Then there exists a net $\langle y_\alpha \rangle \subset A$ such that $y_\alpha \to \phi^{-1}(x)$. Then $\langle \phi(y_\alpha) \rangle \subset \phi(A)$ and $\phi(y_\alpha) \to x$. Thus $x \in \overline{\phi(A)}$ and $\phi(\overline{A}) \subset \overline{\phi(A)}$.
2. Similar

□

3.1. Semi-continuity.

Definition 3.1.1. Let X be a topological space, $f : X \to (\infty, \infty]$ and $x_0 \in X$. Then f is said to be **lower semicontinuous (l.s.c.) at** x_0 if

$$\liminf_{x \to x_0} f(x) \geq f(x_0)$$

and f is said to be **lower semicontinuous (l.s.c.)** if for each $x_0 \in X$, f is lower semicontinuous at x_0.

Exercise 3.1.2. Let X be a topological space and $f : X \to (\infty, \infty]$. Then f is l.s.c. iff for each $\alpha \in \mathbb{R}$, $f^{-1}((\alpha, \infty])$ is open.

Proof. Suppose that f is l.s.c. Let $\alpha \in \mathbb{R}$ and $x_0 \in f^{-1}(\alpha, \infty]$. Put $\epsilon = f(x_0) - \alpha$. By definition,

$$\sup_{V \in N_{x_0}} \inf_{x \in V \setminus \{x_0\}} f(x) \geq f(x_0)$$

Choose $V_\epsilon \in N_{x_0}$ such that

$$\begin{aligned} \inf_{x \in V_\epsilon} f(x) &> f(x_0) - \epsilon \\ &= \alpha \end{aligned}$$

Then $V_\epsilon^o \in \mathcal{N}_{x_0}$ is open and

$$\begin{aligned} V_\epsilon^o &\subset V_\epsilon \\ &\subset f^{-1}((\alpha, \infty]) \end{aligned}$$

So $f^{-1}((\alpha, \infty])$ is open.
Conversely, suppose that for each $\alpha \in \mathbb{R}$, $f^{-1}((\alpha, \infty])$ is open. Let $x_0 \in X$. Put $\alpha = f(x_0)$. For $n \in \mathbb{N}$, define $V_n = f^{-1}((f(x_0) - 1/n, \infty])$. Then for each $n \in \mathbb{N}$, $V_n \in \mathcal{N}_{x_0}$ and

$$\begin{aligned} \liminf_{x \to x_0} f(x) &= \sup_{V \in \mathcal{N}_{x_0}} \inf_{x \in V \setminus \{x_0\}} f(x) \\ &\geq \sup_{n \in \mathbb{N}} \inf_{x \in V_n \setminus \{x_0\}} f(x) \\ &\geq \sup_{n \in \mathbb{N}} f(x_0) - 1/n \\ &= f(x_0) \end{aligned}$$

So f is l.s.c. $\square$

4. Banach Spaces

4.1. **Introduction.**

Note 4.1.1. In the following, we will consider vector spaces over $\mathbb{C}$. There are analogous results for real vector spaces as well, just replace every $\mathbb{C}$ with $\mathbb{R}$.

Definition 4.1.1. Let X be a normed vector space. Then X is said to be a **Banach space** if X is complete.

Definition 4.1.2. Let X be a normed vector space and $(x_i)_{i=1}^n \subset X$. The series $\sum_{i=1}^\infty x_i$ is said to **converge** if the sequence $s_n := \sum_{i=1}^n x_i$ converges. The series $\sum_{i=1}^\infty x_i$ is said to **converge absolutely** if $\sum_{i\in\mathbb{N}} \|x_i\| < \infty$.

Exercise 4.1.3. Let X be a normed vector space. Then X is complete iff for each $(x_i)_{i\in\mathbb{N}} \subset X$, $\sum_{i=1}^\infty x_i$ converges absolutely implies that $\sum_{i=1}^\infty x_i$ converges.
Hint: Given a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$, obtain a subsequence $(x_{n_j})_{j\in\mathbb{N}} \subset (x_n)_{n\in\mathbb{N}}$ such that for each $j \in \mathbb{N}$, $\|x_{n_{j+1}} - x_{n_j}\| < 2^{-j}$. Define a new sequence $(y_j)_{j\in\mathbb{N}} \subset X$ by

$$y_j = \begin{cases} x_{n_1} & j = 1 \\ x_{n_j} - x_{n_{j-1}} & j \geq 2 \end{cases}$$

Proof. Suppose that X is complete. Let $(x_i)_{i\in\mathbb{N}} \subset X$. Suppose that $\sum_{i=1}^\infty x_i$ converges absolutely. Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that for each $m, n \in \mathbb{N}$, if $m, n \geq N$ and $m < n$, then $\sum_{m+1}^n \|x_i\| < \epsilon$. Let $m, n \in \mathbb{N}$. Suppose that $m < n$. Then

$$\begin{aligned} \|s_n - s_m\| &= \left\| \sum_{i=1}^n x_i - \sum_{i=1}^m x_i \right\| \\ &= \left\| \sum_{i=m+1}^n x_i \right\| \\ &\leq \sum_{i=m+1}^n \|x_i\| \\ &< \epsilon \end{aligned}$$

Thus $(s_n)_{n\in N}$ is Cauchy. Since X is complete, $\sum_{i=1}^\infty x_i$ converges.
Conversely, Suppose that for each $(x_i)_{i\in\mathbb{N}} \subset X$, $\sum_{i=1}^\infty x_i$ converges absolutely implies that $\sum_{i=1}^\infty x_i$ converges. Let $(x_i)_{i\in\mathbb{N}} \subset X$ be Cauchy. Proceed inductively to create a strictly increasing sequence $(n_i)_{i\in\mathbb{N}} \subset \mathbb{N}$ such that for each $m, n \in \mathbb{N}$, if $m, n \geq n_i$, then $\|x_m - x_n\| < 2^{-i}$. Define $(y_i)_{i\in\mathbb{N}} \subset X$ by

$$y_i = \begin{cases} x_{n_1} & i = 1 \\ x_{n_i} - x_{n_{i-1}} & i \geq 2 \end{cases}$$

Then $\sum_{i=1}^k y_i = x_{n_k}$ and

$$\begin{aligned} \sum_{i\in\mathbb{N}} \|y_i\| &= \|x_{n_1}\| + \sum_{i\in\mathbb{N}} \|x_{n_i} - x_{n_{i-1}}\| \\ &\leq \|x_{n_1}\| + 2\sum_{i\in\mathbb{N}} 2^{-i} \\ &= \|x_{n_1}\| + 2 \end{aligned}$$

Hence $(x_{n_k})_{k\in\mathbb{N}} = (\sum_{i=1}^{k} y_i)_{i\in\mathbb{N}}$ converges. Since $(x_i)_{i\in\mathbb{N}}$ is cauchy and has a convergent subsequence, it converges. So X is complete. $\square$

4.2. **Linear and Multilinear Maps.**

Definition 4.2.1. Let X, Y be a normed vector spaces. A linear map $T : X \to Y$ is said to be **bounded** if there exists $C \geq 0$ such that for each $x \in X$,

$$\|Tx\| \leq C\|x\|$$

We define

$$L(X,Y) = \{T : X \to Y : T \text{ is linear and bounded}\}$$

Exercise 4.2.2. Let X, Y be a normed vector spaces and $T : X \to Y$ a linear map. Then T is bounded iff there exists $r, s > 0$ such that $T(B(0,r)) \subset B(0,s)$

Proof. Suppose that T is bounded. Then there exists $C \geq 0$ such that for each $x \in X$, $\|Tx\| \leq C\|x\|$. Thus $T(B(0,1)) \subset B(0, C+1)$. Conversely. Suppose that there exists $r, s > 0$ such that $T(B(0,r)) \subset B(0,s)$. Define $C = \frac{2s}{r}$. Let $x \in X$. Put $\alpha = \frac{r}{2\|x\|}$ Then $\alpha x \in B(0,r)$. So $T(\alpha x) = \alpha T(x) \in B(0,s)$. Hence

$$\begin{aligned}
\|T(\alpha x)\| &= \|\alpha T(x)\| \\
&= |\alpha|\|T(x)\| \\
&= \frac{r}{2\|x\|}\|T(x)\| \\
&< s.
\end{aligned}$$

Thus

$$\|Tx\| < \frac{2s}{r}\|x\| = C\|x\|$$

So T is bounded. □

Theorem 4.2.1. *Let X, Y be normed vector spaces and $T : X \to Y$ a linear map. Then the following are equivalent:*

(1) T is continuous
(2) T is continuous at $x = 0$
(3) T is bounded

Proof.

- (1) $\implies$ (2):
  Trivial
- (2) $\implies$ (3):
  Suppose that T is continuous at $x = 0$. Then there exists $\delta > 0$ such that for each $x \in X$, if $\|x\| < \delta$, then $\|Tx\| < 1$. Choose $C = \frac{2}{\delta}$. If $x = 0$, then $\|Tx\| \leq C\|x\|$. Suppose that $\|x\| \neq 0$. Define $y = \frac{\delta}{2\|x\|}x$. Then $\|y\| < \delta$. So

$$\begin{aligned}
1 &> \|Ty\| \\
&= \frac{\delta}{2\|x\|}\|Tx\|
\end{aligned}$$

  Thus

$$\begin{aligned}
\|Tx\| &< \frac{2}{\delta}\|x\| \\
&= C\|x\|
\end{aligned}$$

  Hence T is bounded.

- (3) $\implies$ (1)
  Suppose that T is bounded. Then there exists $C \geq 0$ such that for each $x \in X$, $\|Tx\| \leq C\|x\|$. Let $\epsilon > 0$. Choose $\delta = \frac{\epsilon}{C+1}$. Let $x, y \in X$ Suppose that $\|x - y\| < \delta$. Then
$$\begin{aligned}\|Tx - Ty\| &= \|T(x-y)\| \\ &\leq C\|x-y\| \\ &< (C+1)\delta \\ &= \epsilon\end{aligned}$$
  So T is continuous.

$\square$

Definition 4.2.3. Let X, Y be normed vector spaces. Define $\|\cdot\| : L(X,Y) \to [0,\infty)$ by
$$\|T\| = \inf\{C \geq 0 : \text{for each } x \in X,\ \|Tx\| \leq C\|x\|\}$$
We call $\|\cdot\|$ the **operator norm on** $L(X,Y)$

Exercise 4.2.4. Let X, Y be normed vector spaces. If $X \neq \{0\}$, then the operater norm on $L(X,Y)$ is given by:

(1) $\|T\| = \sup\limits_{\|x\|=1} \|Tx\|$
(2) $\|T\| = \sup\limits_{x\neq 0} \|x\|^{-1}\|Tx\|$
(3) $\|T\| = \inf\{C \geq 0 : \text{for each } x \in X,\ \|Tx\| \leq C\|x\|\}$

Proof. Since $X \neq \{0\}$, the supremums in (1) and (2) are well defined. Let $T \in L(X,Y)$. By linearity of T, the sets over which the supremums are taken in (1) and (2) are the same. So (1) and (2) are equal.

Now, put $M = \sup\limits_{\|x\|=1} \|Tx\|$, $m = \inf\{C \geq 0 : \text{ for each } x \in X,\ \|Tx\| \leq C\|x\|\}$ and let $x \in X$. If $\|x\| = 0$, then $\|Tx\| \leq M\|x\|$. Suppose that $\|x\| \neq 0$. Then
$$\begin{aligned}\|Tx\| &= \bigg(\Big\|T(x/\|x\|)\Big\|\bigg)\|x\| \\ &\leq M||x||\end{aligned}$$
Hence $M \in \{C \geq 0 : \text{ for each } x \in X,\ \|Tx\| \leq C\|x\|\}$. Therefore $m \leq M$

Let $C \in \{C \geq 0 : \text{ for each } x \in X,\ \|Tx\| \leq C\|x\|\}$. Suppose that $\|x\| = 1$. Then $\|Tx\| \leq C\|x\| = C$. So $M \leq C$. Therefore $M \leq m$. So $M = m$ and the supremum in (1) is the same as the infimum in (3). $\square$

Note 4.2.1. From here on, unless stated otherwise, we assume $X \neq 0$.

Exercise 4.2.5. Let X, Y be normed vector spaces and $T \in L(X,Y)$. Then for each $x \in X$, $\|Tx\| \leq \|T\|\|x\|$

Proof. This is just part of the previous exercise. Let $x \in X$. If $x = 0$, then $\|Tx\| \leq \|T\|\|x\|$. Suppose that $x \neq 0$. Then $\|Tx\| = T(x/\|x\|)\|x\| \leq \|T\|\|x\|$ $\square$

Exercise 4.2.6. Let X, Y be normed vector spaces. Then the operator norm is a norm on $L(X, Y)$.

Proof. Let $S, T \in L(X, Y)$ and $\alpha \in \mathbb{C}$. For each $x \in X$, we have that

$$\begin{aligned} \|(S+T)x\| &= \|Sx + Tx\| \\ &\leq \|Sx\| + \|Tx\| \\ &\leq \|S\|\|x\| + \|T\|\|x\| \\ &= \big(\|S\| + \|T\|\big)\|x\| \end{aligned}$$

So $\|S + T\| \leq \|S\| + \|T\|$.

Using the definition of $\|T\|$, we see that

$$\begin{aligned} \|\alpha T\| &= \sup_{\|x\|=1} \|(\alpha T)x\| \\ &= \sup_{\|x\|=1} |\alpha|\|Tx\| \\ &= |\alpha| \sup_{\|x\|=1} \|Tx\| \\ &= |\alpha|\|T\| \end{aligned}$$

So $\|\alpha S\| = |\alpha|\|S\|$.

Suppose that $\|T\| = 0$. Let $x \in X$. Then $\|Tx\| \leq \|T\|\|x\| = 0$. So $Tx = 0$. Since $x \in X$ is arbitrary, we have that $T = 0$. □

Exercise 4.2.7. Let X, Y, Z be normed vector spaces, $T \in L(X, Y)$ and $S \in L(Y, Z)$. Define $ST : X \to Z$ by $STx = S(Tx)$. Then $ST \in L(X, Z)$ and $\|ST\| \leq \|S\|\|T\|$.

Proof. Clearly ST is linear. Let $x \in X$. Then

$$\begin{aligned} \|STx\| &= \|S(Tx)\| \\ &\leq \|S\|\|Tx\| \\ &\leq \|S\|\|T\|\|x\| \end{aligned}$$

So $\|ST\| \leq \|S\|\|T\|$. □

Definition 4.2.8. Let X, Y be a normed vector spaces and $T \in L(X, Y)$. Then T is said to be **invertible** or an **isomorphism** if T is a bijection and $T^{-1} \in L(Y, X)$.

Definition 4.2.9. Let X be a normed vector space. Define $GL(X) := \{T \in L(X, X) : T \text{ is invertible}\}$.

Exercise 4.2.10. Let X be a normed vector space. Then addition and scalar multiplication are continuous on $X \times X$ and $\|\cdot\| : X \to [0, \infty)$ is continuous.

Proof. Let $\epsilon > 0$. Choose $\delta = \frac{\epsilon}{2}$. Let $(x_1, y_1), (x_2, y_2) \in X \times X$. Suppose that $\|(x_1, y_1) - (x_2, y_2)\| = \max\{\|x_1 - x_2\|, \|y_1 - y_2\|\} < \delta$. Then

$$\begin{aligned}\|(x_1 + y_1) - (x_2 + y_2)\| &= \|(x_1 - x_2) + (y_1 - y_2)\| \\ &\leq \|x_1 - x_2\| + \|y_1 - y_2\| \\ &< 2\delta \\ &= \epsilon\end{aligned}$$

Hence addition is uniformly continuous.

Let $(\lambda_1, x_1) \in \mathbb{C} \times X$ and $\epsilon > 0$. Choose $\delta = \min\{\frac{\epsilon}{2(|\lambda_1| + \|x_1\| + 1)}, \frac{\sqrt{\epsilon}}{\sqrt{2}}\}$. Let $(\lambda_2, x_2) \in \mathbb{C} \times X$. Suppose that $\|(\lambda_1, x_1) - (\lambda_2, x_2)\| = \max\{|\lambda_1 - \lambda_2|, \|x_1 - x_2\|\} < \delta$. Then

$$\begin{aligned}\|\lambda_1 x_1 - \lambda_2 x_2\| &= \|\lambda_1 x_1 - \lambda_1 x_2 + \lambda_1 x_2 - \lambda_2 x_2\| \\ &= \|\lambda_1(x_1 - x_2) + (\lambda_1 - \lambda_2)x_2\| \\ &\leq |\lambda_1|\|x_1 - x_2\| + |\lambda_1 - \lambda_2|\|x_2\| \\ &\leq |\lambda_1|\|x_1 - x_2\| + |\lambda_1 - \lambda_2|(\|x_1 - x_2\| + \|x_1\|) \\ &< |\lambda_1|\delta + \delta(\delta + \|x_1\|) \\ &= (|\lambda_1| + \|x_1\|)\delta + \delta^2 \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon\end{aligned}$$

Since $(\lambda_1, x_1) \in \mathbb{C} \times X$ is arbitrary, scalar multiplication is continuous.

Let $\epsilon > 0$. Choose $\delta = \epsilon$. Let $x, y \in X$. Suppose that $\|x - y\| < \delta$. Then

$$\begin{aligned}\big|\|x\| - \|y\|\big| &\leq \|x - y\| \\ &< \delta \\ &= \epsilon\end{aligned}$$

So $\|\cdot\| : X \to [0, \infty)$ is uniformly continuous. $\square$

Exercise 4.2.11. Let X, Y be normed vector spaces. If Y is complete, then so is $L(X, Y)$.

Proof. Suppose that Y is complete. Let $(T_n)_{n\in\mathbb{N}} \subset L(X, Y)$. Suppose that $(T_n)_{n\in\mathbb{N}}$ is Cauchy. Since for each $m, n \in \mathbb{N}$, $\big|\|T_m\| - \|T_n\|\big| \leq \|T_m - T_n\|$, we have that $(\|T_n\|)_{n\in\mathbb{N}} \subset [0, \infty)$ is Cauchy. Hence $\lim\limits_{n\to\infty} \|T_n\|$ exists.

Let $x \in X$ and $m, n \in \mathbb{N}$. Then

$$\begin{aligned}\|T_m x - T_n x\| &= \|(T_m - T_n)x\| \\ &\leq \|T_m - T_n\|\|x\|\end{aligned}$$

So $(T_n x)_{n\in\mathbb{N}} \subset Y$ is Cauchy and hence converges. Define $T : X \to Y$ by $Tx = \lim\limits_{n\to\infty} T_n x$.

Since addition and scalar multiplication are continuous, T is linear. Let $x \in X$ and $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that for each $n \in N$, if $n \geq N$, then $\|Tx - T_n x\| < \epsilon$. Then for each $n \in \mathbb{N}$, if $n \geq N$ we have that

$$\begin{aligned} \|Tx\| &\leq \|Tx - T_n x\| + \|T_n x\| \\ &< \epsilon + \|T_n x\| \\ &\leq \epsilon + \|T_n\| \|x\| \end{aligned}$$

Thus $\|Tx\| \leq \epsilon + (\lim_{n \to \infty} \|T_n\|)\|x\|$. Since $\epsilon > 0$ is arbitrary, $\|Tx\| \leq (\lim_{n \to \infty} \|T_n\|)\|x\|$. Thus $T \in L(X, Y)$ and $\|T\| \leq \lim_{n \to \infty} \|T_n\|$.

Note that since addition, scalar multiplication and $\|\cdot\|$ are continuous, we have that for each $n \in \mathbb{N}$ and $x \in X$, $\|(T_n - T_m)x\|$ converges to $\|(T_n - T)x\|$ because

$$\begin{aligned} \lim_{m \to \infty} \|(T_n - T_m)x\| &= \lim_{m \to \infty} \|T_n x - T_m x\| \\ &= \|T_n x - \lim_{m \to \infty} T_m x\| \\ &= \|T_n x - Tx\| \\ &= \|(T_n - T)x\| \end{aligned}$$

Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that for each $m, n \in \mathbb{N}$ if $n, m \geq N$, then $\|T_n - T_m\| < \epsilon$. Then for each $n \in \mathbb{N}$ if $n \geq N$, then for each $x \in X$,

$$\|(T_n - T_m)x\| \leq \|(T_n - T_m)\| \|x\| < \epsilon \|x\|$$

Combining this with the previous fact, we see that for each $n \in N$, if $n \geq N$, then for each $x \in X$,

$$\|(T_n - T)x\| \leq \epsilon \|x\|$$

In particular, for each $n \in \mathbb{N}$, if $n \geq N$, then

$$\|T_n - T\| = \sup_{\|x\|=1} \|(T_n - T)x\| \leq \epsilon$$

This implies that T_n converges to T in $L(X, Y)$. Since

$$\big| \|T_n\| - \|T\| \big| \leq \|T_n - T\|$$

it is clear that $\lim_{n \to \infty} \|T_n\| = \|T\|$ $\square$

Definition 4.2.12. Let $X_1, \cdots, X_n, Y$ be normed vector spaces and $T : \prod_{i=1}^{n} X_i \to Y$ multilinear. Then T is said to be **bounded** if there exists $C \geq 0$ such that for each $x_1, \cdots, x_n \in X$,

$$\|T(x_1, \cdots, x_n)\| \leq C \|x_1\| \cdots \|x_n\|$$

We define

$$L^n(X_1, \ldots, X_n; Y) = \left\{ T : \prod_{i=1}^{n} X_i \to Y : T \text{ is multilinear and bounded} \right\}$$

If $X_1, \cdots, X_n = X$, we write $L^n(X, Y)$ in place of $L^n(X, \ldots, X; Y)$

Exercise 4.2.13. Let $X_1, \cdots, X_n, Y$ be normed vector spaces and $T : \prod_{i=1}^{n} X_i \to Y$ multilinear. Then the following are equivalent:

(1)
(2)
(3)

4.3. Quotient Spaces.

Definition 4.3.1. Let X be a normed vector space and $M \subset X$ a closed subspace. Define $\|\cdot\| : X/M \to [0,\infty)$ by

$$\|x+M\| := \inf_{y\in M} \|x+y\|$$

We call $\|\cdot\|$ the **subspace norm on** X/M

Exercise 4.3.2. Let X be a normed vector space and $M \subsetneq X$ a proper, closed subspace of M. Then

(1) The previously defined subspace norm on X/M is well defined and is a norm.
(2) For each $\epsilon > 0$, there exists $x \in X$ such that $\|x\| = 1$ and $\|x+M\| \geq 1-\epsilon$.
(3) The projection map $\pi : X \to X/M$ defined by $\pi(x) = x+M$ is continuous and $\|\pi\| = 1$.
(4) If X is complete, then X/M is complete.

Proof. (1) Let $x, y \in X$ and $\alpha \in \mathbb{C}$. Suppose that $x+M = y+M$. Then there exists $m \in M$ such that $x = y+m$. Since M is a subspace, the map $T : M \to M$ given by $Tx = x+m$ is a bijection. So

$$\inf_{z\in M} \|y+m+z\| = \inf_{z\in M} \|y+z\|$$

which implies that

$$\begin{aligned} \|x+M\| &= \inf_{z\in M} \|x+z\| \\ &= \inf_{z\in M} \|y+m+z\| \\ &= \inf_{z\in M} \|y+z\| \\ &= \|y+M\| \end{aligned}$$

So $\|\cdot\| : X/M \to [0,\infty)$ is well defined.

We observe that for each $z, w \in M$,

$$\|x+y+z\| \leq \|x+w\| + \|y+w+z\|$$

Taking infimums over M with respect to z in this inequality implies that for each $w \in M$,

$$\begin{aligned} \inf_{z\in M} \|x+y+z\| &\leq \inf_{z\in M} \Big(\|x+w\| + \|y+w+z\|\Big) \\ &= \|x+w\| + \inf_{z\in M} \|y+w+z\| \end{aligned}$$

Again we use the fact that for each $w \in M$,

$$\inf_{z\in M} \|y+w+z\| = \inf_{z\in M} \|y+z\|$$

This implies that for each $w \in M$,

$$\inf_{z\in M} \|x+y+z\| \leq \|x+w\| + \inf_{z\in M} \|y+z\|$$

Therefore, taking infimums over M with respect to w in this inequality yields

$$\begin{aligned}\|x+y+M\| &= \inf_{z\in M}\|x+y+z\| \\ &\leq \inf_{w\in M}\left(\|x+w\|+\inf_{z\in M}\|y+z\|\right) \\ &= \inf_{w\in M}\|x+w\|+\inf_{z\in M}\|y+z\| \\ &= \|x+M\|+\|y+M\|\end{aligned}$$

If $\alpha=0$, then $\alpha x=0$. Choosing $z=0\in M$ gives $\|\alpha x+M\|=0=|\alpha|\|x+M\|$. Suppose that $\alpha\neq 0$. Then the map $T:M\to M$ given by $Tx=\alpha^{-1}x$ is a bijection and thus $\inf\limits_{z\in M}\|x+\alpha^{-1}z\|=\inf\limits_{z\in M}\|x+z\|$. Hence we have that

$$\begin{aligned}\|\alpha x+M\| &= \inf_{z\in M}\|\alpha x+z\| \\ &= \inf_{z\in M}|\alpha|\|x+\alpha^{-1}z\| \\ &= |\alpha|\inf_{z\in M}\|x+\alpha^{-1}z\| \\ &= |\alpha|\inf_{z\in M}\|x+z\| \\ &= |\alpha|\|x+M\|\end{aligned}$$

Suppose that $\|x\|=0$. Choose a sequence $(z_n)_{n\in N}\subset M$ such that

$$\begin{aligned}\lim_{n\to\infty}\|x-z_n\| &= \inf_{z\in M}\|x+z\| \\ &= 0\end{aligned}$$

Then $\lim\limits_{n\to\infty}z_n=x$. Since M is closed, $x\in M$. Hence $x+M=0+M$.

(2) Since M is a proper subspace, there exists $v\in X$ such that $v\notin M$. Then $\|v+M\|\neq 0$. Let $\epsilon>0$. Then $(1-\epsilon)^{-1}\|v+M\|>\|v+M\|$. So there exists $z\in M$ such that

$$0<\|v+M\|\leq\|v+z\|<(1-\epsilon)^{-1}\|v+M\|$$

Choose $x=\|v+z\|^{-1}(v+z)$. Then $\|x\|=1$ and

$$\begin{aligned}\|x+M\| &= \|v+z\|^{-1}\|v+z+M\| \\ &= \|v+z\|^{-1}\|v+M\| \\ &> 1-\epsilon\end{aligned}$$

(3) Let $x\in X$. Taking $z=0$, we we see that $\|\pi(x)\|=\|x+M\|\leq\|x+z\|=\|x\|$. So π is bounded and in particular,

$$\sup_{\|x\|=1}\|\pi(x)\|\leq 1$$

From (2) we see that

$$\sup_{\|x\|=1} \|\pi(x)\| \geq 1$$

Hence $\|\pi\| = 1$.

(4) Suppose that X is complete. Let $(x_i + M)_{i\in\mathbb{N}} \subset X/M$. Suppose that $\sum\limits_{i\in\mathbb{N}} \|x_i + M\| < \infty$. Let $\epsilon > 0$. Then for each $i \in \mathbb{N}$, there exists $z_i \in M$ such that $\|x_i + z_i\| < \|x_i + M\| + \epsilon 2^{-i}$. Define the sequence $(a_i)_{i\in\mathbb{N}} \subset X$ by $a_i = x_i + z_i$. Then we have

$$\begin{aligned}
\sum_{i\in\mathbb{N}} \|a_i\| &= \sum_{i\in N} \|x_i + z_i\| \\
&\leq \sum_{i\in N} \left(\|x_i + M\| + \epsilon 2^{-i}\right) \\
&= \sum_{i\in\mathbb{N}} \|x_i + M\| + \epsilon
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that

$$\sum_{i\in\mathbb{N}} \|a_i\| \leq \sum_{i\in\mathbb{N}} \|x_i + M\| < \infty$$

Since X is complete, $\sum\limits_{i=1}^{\infty} a_i$ converges in X. Define $(s_n)_{n\in\mathbb{N}} \subset X$ and $s \in X$ by $s_n = \sum\limits_{i=1}^{n} a_i$ and $s = \sum\limits_{i=1}^{\infty} a_i$. Since $\lim\limits_{n\to\infty} s_n = s$, and $\pi : X \to X/M$ is continuous, it follows that $\lim\limits_{n\to\infty} \pi(s_n) = \pi(s)$. Since

$$\begin{aligned}
\pi(s_n) &= \sum_{i=1}^{n} a_i + M \\
&= \sum_{i=1}^{n} x_i + M
\end{aligned}$$

We have that $\sum\limits_{i=1}^{\infty} x_i + M$ converges which implies that X/M is complete.

□

Exercise 4.3.3. Let X, Y be normed vector spaces and $T \in L(X, Y)$. Then

(1) $\ker T$ is closed
(2) there exists a unique map $S : X/\ker T \to T(X)$ such that $T = S \circ \pi$. Furthermore S is a bounded linear bijection and $\|S\| = \|T\|$.

Proof. (1) Since T is continuous and $\ker T = T^{-1}(\{0\})$, we have that $\ker T$ is closed.

(2) Suppose that there exists $S_1, S_2 \in L(X/\ker T, T(X))$ such that $T = S_1 \circ \pi$ and $T = S_2 \circ \pi$. Let $x \in X$. Then

$$S_1(x + \ker T) = S_1(\pi(x)) = T(x) = S_2(\pi(x)) = S_2(x + \ker T)$$

So $S_1 = S_2$. Therefore such a map is unique.
Define $S : X/\ker T \to T(X)$ by $S(x + \ker T) = T(x)$. Then S is clearly a linear bijection that satisfies $T = S \circ \pi$. Let $x \in X$ and $z \in \ker T$. Then

$$\begin{aligned}\|S(x + \ker T)\| &= \|T(x)\| \\ &= \|T(x + z)\| \\ &\leq \|T\|\|x + z\|\end{aligned}$$

Thus

$$\|S(x + \ker T)\| \leq \|T\| \inf_{z \in \ker T} \|x + z\| = \|T\|\|x + \ker T\|$$

So S is bounded and $\|S\| \leq \|T\|$. This implies that

$$\|T\| = \|S \circ \pi\| \leq \|S\|\|\pi\| = \|S\|$$

Thus $\|S\| = \|T\|$.

□

Exercise 4.3.4. Let X, Y be normed vector spaces. Define $\phi : L(X, Y) \times X \to Y$ by $\phi(T, x) = Tx$. Then ϕ is continuous.

Proof. Let $(T_1, x_1) \in L(X, Y) \times X$ and $\epsilon > 0$. Choose $\delta = \min\{\frac{\epsilon}{2(\|x_1\| + \|T_1\| + 1)}, \frac{\sqrt{\epsilon}}{\sqrt{2}}\}$. Let $(t_2, x_2) \in L(X, Y) \times X$. Suppose that

$$\|(T_1, x_1) - (T_2, x_2)\| = \max\{\|T_1 - T_2\|, \|x_1 - x_2\|\} < \delta$$

Then

$$\begin{aligned}\|\phi(T_1, x_1) - \phi(T_2 - x_2)\| &= \|T_1 x_- T_2 x_2\| \\ &= \|T_1 x_1 - T_2 x_1 + T_2 x_1 - T_2 x_2\| \\ &\leq \|(T_1 - T_2) x_1\| + \|T_2(x_1 - x_2)\| \\ &\leq \|T_1 - T_2\|\|x_1\| + \|T_2\|\|x_1 - x_2\| \\ &\leq \|T_1 - T_2\|\|x_1\| + \big(\|T_1 - T_2\| + \|T_1\|\big)\|x_1 - x_2\| \\ &< \delta\|x_1\| + (\delta + \|T_1\|)\delta \\ &= \delta(\|T_1\| + \|x_1\|) + \delta^2 \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon\end{aligned}$$

So ϕ is continuous. □

Exercise 4.3.5. Let X be a normed vector space and $M \subset X$ a subspace. Then $\overline{M}$ is a subspace.

Proof. Let $x, y \in \overline{M}$ and $\alpha \in \mathbb{C}$. Then there exist sequences $(x_n)_{n \in \mathbb{N}} \subset M$ and $(y_n)_{n \in \mathbb{N}} \subset M$ such that $x_n \to x$ and $y_n \to y$. Since M is a subspace, $(x_n + y_n)_{n \in \mathbb{N}} \subset M$ and $(\alpha x_n)_{n \in \mathbb{N}} \subset M$. Since addition and scalar multiplication are continuous, we have that $x_n + y_n \to x + y$ and $\alpha x_n \to \alpha x$. Thus $x + y \in \overline{M}$ and $\alpha x \in \overline{M}$ and hence $\overline{M}$ is a subspace. □

4.4. **Direct Sums.**

Definition 4.4.1. Let $X_1, \cdots, X_n$ be Banach spaces and $p \in [1, \infty]$. We define $\| \cdot \|_p : \bigoplus_{j=1}^{n} X_j \to [0, \infty)$ by

$$\|(x_j)_{j=1}^n\|_p = \|(\|x_j\|)_{j=1}^n\|_{\mathbb{R}^n, p}$$

Exercise 4.4.2. Let $X_1, \cdots, X_n$ be Banach spaces. Then

(1) for each $p \in [1, \infty]$, $\| \cdot \|_p : \bigoplus_{j=1}^{n} X_j \to [0, \infty)$ is a norm on $\bigoplus_{j=1}^{n} X_j$
(2) $\{\| \cdot \|_p : p \in [1, \infty]\}$ are equivalent.

Proof.

(1) Let $p \in [1, \infty]$.
(2)

$\square$

4.5. **Linear and Sublinear Functionals.**

Definition 4.5.1.

(1) Let X be a vector space over $\mathbb{C}$ and $T : X \to \mathbb{C}$. Then T is said to be a **linear functional on** X if T is linear. We define the **dual space of** X, denoted X^*, by $X^* = \{T : X \to \mathbb{C} : T \text{ is linear}\}$

(2) If X is a normed vector space over $\mathbb{C}$, then T is said to be a **bounded linear functional on** X if $T \in L(X, \mathbb{C})$. We define the **dual space of** X, denoted X^*, by $X^* = L(X, \mathbb{C})$.

Note 4.5.1. We define X^* similarly when X is an vector space or normed vector space over $\mathbb{R}$.

Definition 4.5.2. Let X be a normed vector space and $p : X \to \mathbb{R}$. Then p is said to be a **sublinear functional** if for each $x, y \in X$, $\lambda \geq 0$,

(1) $p(x + y) \leq p(x) + p(y)$

(2) $p(\lambda x) = \lambda p(x)$

Exercise 4.5.3. Let X be a vector space and $\|\cdot\| : X \to [0, \infty)$ be a seminorm, then $\|\cdot\|$ is a sublinear functional.

Proof. Clear □

Exercise 4.5.4. Let X be a vector space and $p : X \to \mathbb{R}$ a sublinear functional. Then for each $x, y \in X$

(1) $-p(-x) \leq p(x)$

(2) $-p(y - x) \leq p(x) - p(y) \leq p(x - y)$

Proof. Let $x, y \in X$.

(1) We have

$$\begin{aligned} 0 &= p(0) \\ &= p(x - x) \\ &\leq p(x) + p(-x) \end{aligned}$$

So $-p(-x) \leq p(x)$.

(2) We have

$$\begin{aligned} p(x) &= p(x - y + y) \\ &\leq p(x - y) + p(y) \end{aligned}$$

So $p(x) - p(y) \leq p(x - y)$. Switching x and y gives us $p(y) - p(x) \leq p(y - x)$ and multiplying both sides by -1 yields $-p(y - x) \leq p(x) - p(y)$

Putting these two together, we see that

$$-p(y - x) \leq p(x) - p(y) \leq p(x - y)$$

□

Definition 4.5.5. Let X be a normed vector space and $p : X \to \mathbb{R}$ a sublinear functional. Then p is said to be **bounded** if there exists $M > 0$ such that for each $x \in X$, $p(x) \leq M\|x\|$.

Exercise 4.5.6. Let X be a normed vector space and $p : X \to \mathbb{R}$ a sublinear functional. Then p is bounded iff p is Lipschitz.

Proof. Suppose that p is bounded. Then there exists $M > 0$ such that for each $x \in X$, $p(x) \leq M\|x\|$. Let $x, y \in X$. Then the previous exercise implies that

$$\begin{aligned} -M\|x-y\| &= -M\|y-x\| \\ &\leq -p(y-x) \\ &\leq p(x)-p(y) \\ &\leq p(x-y) \\ &\leq M\|x-y\| \end{aligned}$$

So that

$$|p(x)-p(y)| \leq M\|x-y\|$$

and p is Lipschitz. Conversely, suppose that p is Lipschitz. Then there exists $M > 0$ such that for each $x, y \in X$, $|p(x)-p(y)| \leq M\|x-y\|$. Let $x \in X$. Then

$$\begin{aligned} p(x) &\leq |p(x)| \\ &= |p(x)-p(0)| \\ &\leq M\|x-0\| \\ &\leq M\|x\| \end{aligned}$$

So p is bounded. □

Theorem 4.5.1. ***Hahn-Banach Theorem:*** *Let X be a vector space, $p : X \to \mathbb{R}$ a sublinear functional, $M \subset X$ a subspace and $f : M \to \mathbb{R}$ a linear functional. If for each $x \in M$, $f(x) \leq p(x)$, then there exists a linear functional $F : X \to \mathbb{R}$ such that for each $x \in X$, $F(x) \leq p(x)$ and $F|_M = f$.*

Exercise 4.5.7. Let X be a vector space and $p : X \to \mathbb{R}$ a sublinear functional. Then there exists $F : X \to \mathbb{R}$ such that for each $x \in X$, $F(x) \leq p(x)$.

Proof. Take $M = \{0\}$ and $f \equiv 0$ and apply the Hahn-Banach theorem. □

Exercise 4.5.8. Equivalency of linearity (General Case) Let X be a vector space and $p : X \to \mathbb{R}$ a sublinear functional. Then the following are equivalent:

(1) there exists a unique $F \in X^*$ such that $F \leq p$
(2) for each $x \in X$, $-p(-x) = p(x)$
(3) p is linear

Hint: If there exists $x \in X$ such that $-p(-x) \neq p(x)$, define $f_1, f_2 : \mathrm{span}(x) \to \mathbb{R}$ by $f_1(tx) = tp(x)$ and $f_2(tx) = -tp(-x)$

Proof.

- (1) $\Rightarrow$ (2):
  Suppose that there exists a unique $F \in X^*$ such that $F \leq p$. For the sake of contradiction, suppose that there exists $x \in X$ such that $-p(-x) \neq p(x)$. Define $f_1, f_2 : \mathrm{span}(x) \to \mathbb{R}$ by

  $$f_1(tx) = tp(x)$$

  and

  $$f_2(tx) = -tp(-x)$$

Let $y \in \text{span}(x)$. Then there exists $t \in \mathbb{R}$ such that $y = tx$. Then for each $k \in \mathbb{R}$,

$$\begin{aligned} f_1(ky) &= f_1(ktx) \\ &= ktp(x) \\ &= kf_1(tx) \\ &= kf_1(y) \end{aligned}$$

Similarly, $f_2(ky) = kf_2(y)$ and so $f_1, f_2 \in \text{span}(x)^*$. If $t \geq 0$, then

$$\begin{aligned} f_1(y) &= f_1(tx) \\ &= tp(x) \\ &= p(tx) \\ &= p(y) \end{aligned}$$

If $t < 0$, then

$$\begin{aligned} f_1(y) &= f_1(tx) \\ &= tp(x) \\ &= -|t|p(x) \\ &= -p(|t|x) \\ &= -p(-tx) \\ &\leq p(tx) \\ &= p(y) \end{aligned}$$

So $f_1 \leq p$ on $\text{span}(x)$. Similarly, $f_2 \leq p$ on $\text{span}(x)$. The Hahn-Banach theorem implies that there exist $F_1, F_2 \in X^*$ such that $F_1, F_2 \leq p$ and $F_1 = f_1, F_2 = f_2$ on $\text{span}(x)$. By the assumption of uniqueness, $F_1 = F_2$. This is a contradiction since

$$\begin{aligned} F_1(x) &= p(x) \\ &\neq -p(-x) \\ &= F_2(x) \end{aligned}$$

So for each $x \in X$, $-p(-x) = p(x)$.

- (2) $\Rightarrow$ (3):
  Suppose that for each $x \in X$, $-p(-x) = p(x)$. The previous exercise implies that there exists $F \in X^*$ such that $F \leq p$. Let $x \in X$. Then

$$\begin{aligned} -F(x) &= F(-x) \\ &\leq p(-x) \\ &= -p(x) \end{aligned}$$

  So $p(x) \leq F(x)$ and $p \leq F$. Therefore $p = F$ and p is linear.
- (3) $\Rightarrow$ (1):
  Suppose that p is linear. Let $F \in X^*$. Suppose that $F \leq p$. Let $x \in X$. Then as in

the case for (2) $\Rightarrow$ (3), we have that

$$\begin{aligned} -F(x) &= F(-x) \\ &\leq p(-x) \\ &= -p(x) \end{aligned}$$

which implies that $p = F$. So p is the unique linear function $F \in X^*$ such that $F \leq p$. $\square$

Exercise 4.5.9. Let X be a normed vector space, $p : X \to \mathbb{R}$ a bounded sublinear functional and $\phi : X \to \mathbb{R}$ a linear functional. If $\phi \leq p$, then $\phi \in X^*$.

Proof. Since p is Lipschitz, there exists $M > 0$ such that for each $x \in X$, $|p(x)| \leq M\|x\|$. Let $x \in X$. Then

$$\begin{aligned} \phi(x) &\leq p(x) \\ &\leq |p(x)| \\ &\leq M\|x\| \end{aligned}$$

and therefore

$$\begin{aligned} -M\|x\| &= -M\|-x\| \\ &\leq -p(-x) \\ &\leq -\phi(-x) \\ &= \phi(x) \end{aligned}$$

So that $|\phi(x)| \leq M\|x\|$ and $\phi \in X^*$. $\square$

Exercise 4.5.10. Let X be a normed vector space and $p : X \to \mathbb{R}$ a bounded sublinear functional. Then there exists $\phi \in X^*$ such that for each $x \in X$, $\phi(x) \leq p(x)$.

Proof. A previous exercise implies there exists $\phi : X \to \mathbb{R}$ such that ϕ is linear and $\phi \leq p$. The previous exercise implies that $\phi \in X^*$. $\square$

Exercise 4.5.11. Equivalency of linearity (Bounded Case) Let X be a normed vector space and $p : X \to \mathbb{R}$ a bounded sublinear functional. Then the following are equivalent:

(1) there exists a unique $\phi \in X^*$ such that $\phi \leq p$
(2) for each $x \in X$, $-p(-x) = p(x)$
(3) p is linear

Proof. Basically the same as last time. $\square$

Theorem 4.5.2. ***Complex Hahn-Banach Theorem:*** *Let X be a vector space, $p : X \to \mathbb{R}$ a seminorm, $M \subset X$ a subspace and $f : M \to \mathbb{C}$ a linear functional. If for each $x \in M$, $|f(x)| \leq p(x)$, then there exists a linear functional $F : X \to \mathbb{C}$ such that for each $x \in X$, $|F(x)| \leq p(x)$ and $F|_M = f$.*

Exercise 4.5.12. Let X be a normed vector space, $M \subset X$ a subspace and $f \in M^*$. Then there exists $F \in X^*$ such that $\|F\| = \|f\|$ and $F|_M = f$.

Proof. If $f = 0$, Choose $F = 0$. Suppose $f \neq 0$. Then $\|f\| \neq 0$ and there exists $x_0 \in M$ such that $x_0 \neq 0$. Thus $\|f\| = \sup\{|f(x)| : x \in M \text{ and } \|x\| = 1\}$. Define $p : X \to [0, \infty)$ by $p(x) = \|f\|\|x\|$. Then p is a sublinear functional on X and for each $x \in M$, $|f(x)| \leq p(x)$. So there exists a linear functional $F : X \to \mathbb{C}$ such that for each $x \in X$, $|F(x)| \leq p(x) = \|f\|\|x\|$ and $F|_M = f$. Thus $F \in X^*$ with $\|F\| \leq \|f\|$. Also

$$\|F\| = \sup_{\substack{x \in X \\ \|x\|=1}} |F(x)| \geq \sup_{\substack{x \in M \\ \|x\|=1}} |F(x)| = \sup_{\substack{x \in M \\ \|x\|=1}} |f(x)| = \|f\|$$

So $\|F\| = \|f\|$. ☐

Exercise 4.5.13. Let X be a normed vector space, $M \subsetneq X$ a proper closed subspace and $x \in X \setminus M$. Then there exists $F \in X^*$ such that $F|_M = 0$, $\|F\| = 1$ and $F(x) = \|x + M\| \neq 0$. **Hint:** Consider $f : M + \mathbb{C}x \to \mathbb{C}$ defined by $f(m + \lambda x) = \lambda \|x + M\|$.

Proof. Define $f : M + \mathbb{C}x \to \mathbb{C}$ as above. Clearly f is linear and $f|M = 0$. Let $m \in M$ and $\lambda \in \mathbb{C}$. If $\lambda = 0$, then $|f(m + \lambda x)| = 0 \leq \|m + \lambda x\|$. Suppose that $\lambda \neq 0$. Then

$$\begin{aligned} |f(m + \lambda x)| &= |\lambda|\|x + M\| \\ &= \|\lambda x + M\| \\ &= \inf_{z \in M} \|z + \lambda x\| \\ &\leq \|m + \lambda x\| \end{aligned}$$

So $f \in (M + \mathbb{C}x)^*$ and $\|f\| \leq 1$. Let $\epsilon > 0$. A previous exercise tells us that there exist $m \in M, \lambda \in \mathbb{C}$ such that $\|m + \lambda x\| = 1$ and $\|m + \lambda x + M\| > 1 - \epsilon$. Then

$$\begin{aligned} |f(m + \lambda x)| &= |\lambda|\|x + M\| \\ &= \|\lambda x + M\| \\ &= \|m + \lambda x + M\| \\ &> 1 - \epsilon \end{aligned}$$

So

$$\|f\| = \sup_{\substack{z \in M + \mathbb{C}x \\ \|z\|=1}} |f(z)| \geq 1$$

Hence $\|f\| = 1$. The same exercise also tells us that $f(x) = \|x + M\| \neq 0$. Using the previous exercise, there exists $F \in X^*$ such that $\|F\| = \|f\| = 1$ and $F|_{M+\mathbb{C}x} = f$. ☐

Exercise 4.5.14. Let X be a normed vector space and $x \in X$. If $x \neq 0$, then there exists $F \in X^*$ such that $\|F\| = 1$ and $F(x) = \|x\|$.

Proof. Define $f : \mathbb{C}x \to \mathbb{C}$ by $f(\lambda x) = \lambda \|x\|$. Then f is linear and $f(x) = \|x\|$. Clearly

$$\sup_{\substack{z \in \mathbb{C}x \\ \|z\|=1}} |f(z)| = 1$$

So $f \in (\mathbb{C}x)^*$ and $\|f\| = 1$. By a previous exercise, there exists $F \in X^*$ such that $\|F\| = \|f\| = 1$ and $F|_{\mathbb{C}x} = f$. ☐

Exercise 4.5.15. Let X be a normed vector space. Then X^* separates the points of X.

Proof. Let $x, y \in X$. Suppose that $x \neq y$. Then $x - y \neq 0$. The previous exercies implies that there exists $F \in X^*$ such that $\|F\| = 1$ and

$$F(x) - F(y) = F(x - y) = \|x - y\| \neq 0$$

Thus $F(x) \neq F(y)$ and X^* separates the points of X. □

Definition 4.5.16. Let X, Y be metric spaces and $T : X \to Y$. Then T is said to be an **isometry** if for each $x_1, x_2 \in X$, $d(Tx_1, Tx_2) = d(x_1, x_2)$.

Exercise 4.5.17. Let X, Y be metric spaces and $T : X \to Y$ and isometry. Then T is injective.

Proof. Let $x_1, x_2 \in X$. Suppose that $Tx_1 = Tx_2$. Then $0 = d(Tx_1, Tx_2) = d(x_1, x_2)$. So $x_1 = x_2$. Hence T is injective. □

Note 4.5.2. Let X, Y be metric spaces and $T : X \to Y$ an isometry. Then T is clearly continuous. If T is surjective, then T^{-1} is an isometry and therefore continuous. Hence T is a homeomorphism.

Exercise 4.5.18. Let X be a normed vector space and $x \in X$. Define $\hat{x} : X^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$. Then $\hat{x} \in X^{**}$ and $\|\hat{x}\| = \|x\|$.

Proof. Let $f, g \in X^*$ and $\lambda \in \mathbb{C}$. Then

$$\hat{x}(f + \lambda g) = (f + \lambda g)(x) = f(x) + \lambda g(x) = \hat{x}(f) + \lambda \hat{x}(g)$$

So $\hat{x}$ is linear. For each $f \in X^*$,

$$|\hat{x}(f)| = |f(x)| \leq \|x\|\|f\|$$

Hence $\hat{x} \in X^{**}$ with $\|\hat{x}\| \leq \|x\|$. If $x = 0$, then $\hat{x} = 0$ and $\|\hat{x}\| = \|x\|$. Suppose that $x \neq 0$. Then a previous exercise implies that there exists $F \in X^*$ such that $\|F\| = 1$ and $F(x) = \|x\|$. Then we have that

$$\sup_{\substack{f \in X^* \\ \|f\|=1}} |\hat{x}(f)| = \sup_{\substack{f \in X^* \\ \|f\|=1}} |f(x)| \geq |F(x)| = \|x\|$$

Hence $\|\hat{x}\| = \|x\|$. □

Exercise 4.5.19. Let X be a normed vector space. Define $\phi : X \to X^{**}$ by $\phi(x) = \hat{x}$. Then ϕ is a linear isometry.

Proof. Let $x, y \in X$ and $\lambda \in \mathbb{C}$. Then for each $f \in X^*$, we have that

$$\begin{aligned}
\phi(x + \lambda y)(f) &= \widehat{x + \lambda y}(f) \\
&= f(x + \lambda y) \\
&= f(x) + \lambda f(y) \\
&= \hat{x}(f) + \lambda \hat{y}(f) \\
&= \phi(x)(f) + \lambda \phi(y)(f)
\end{aligned}$$

So $\phi(x + \lambda y) = \phi(x) + \lambda \phi(y)$ and ϕ is linear. The previous exercise tells us that

$$\begin{aligned}
\|\phi(x) - \phi(y)\| &= \|\phi(x - y)\| \\
&= \|\widehat{x - y}\| = \|x - y\|
\end{aligned}$$

So ϕ is an isometry. □

Definition 4.5.20. Let X be a normed vector space and define $\phi : X \to X^{**}$ as above. We define $\widehat{X} = \phi(X) \subset X^{**}$. Since $\widehat{X}$ and X are isomorphic, we may identify X as a subset of X^{**}.

Definition 4.5.21. Let X be a normed vector space and define $\phi : X \to X^{**}$ as above. Then X is said to be reflexive if ϕ is surjective. In this case ϕ is then an isomorphism

Exercise 4.5.22. Let X be a normed vector space and $f : X \to \mathbb{C}$ a linear functional on X. Then f is bounded iff $\ker f$ is closed.

Proof. Suppose that f is continuous. Since $\{0\}$ is closed, we have that $\ker f = f^{-1}(\{0\})$ is closed. Conversely, suppose that $\ker f$ is closed. If $\ker f = X$, then $f = 0$ and f is continuous. Suppose that $\ker f \neq X$. Then $\ker f$ is a proper, closed subspace of X. A previous exercise tells us that there exists $x \in X$ such that $\|x\| = 1$ and $\|x + \ker f\| > \frac{1}{2}$. Let $y \in X$. Suppose that $\|y\| < \frac{1}{2}$. Then for each $z \in \ker f$,

$$\begin{aligned} \|z - (x + y)\| &= \|(z - x) - y\| \\ &\geq \|z - x\| - \|y\| \\ &> \frac{1}{2} - \frac{1}{2} \\ &= 0 \end{aligned}$$

So $x + y \notin \ker f$. Therefore $f(B(x, \frac{1}{2})) \cap \{0\} = \varnothing$. If $f(B(x, \frac{1}{2}))$ is unbounded, then $f(B(x, \frac{1}{2})) = \mathbb{C}$ by linearity. This is a contradiction since $0 \notin f(B(x, \frac{1}{2}))$. So There exists $s > 0$ such that $f(B(x, \frac{1}{2})) \subset B(0, s)$ and thus f is bounded. $\square$

Exercise 4.5.23. Let X be a normed vector space.

(1) Let $M \subsetneq X$ be a proper closed subspace of X and $x \in X \setminus M$. Then $M + \mathbb{C}x$ is closed.
(2) Let $M \subset X$ be a finite dimensional subspace of X. Then M is closed.

Proof. (1) Let $y \in X$ and $(y_n)_{n \in \mathbb{N}} \subset M + \mathbb{C}x$. Suppose that $y_n \to y$. If $y \in M$, then $y \in M + \mathbb{C}x$. Suppose that $y \notin M$. For each $n \in \mathbb{N}$, there exists $m_n \in M$ and $\lambda_n \in \mathbb{C}$ such that $y_n = m_n + \lambda_n x$. A previous exercise tells us that there exists $F \in X^*$ such that $\|F\| = 1$, $F|_M = 0$ and $F(x) = \|x + M\| \neq 0$. Since F is continuous, $F(y_n) \to F(y)$. Since for each $n \in \mathbb{N}$,

$$F(y_n) = F(m_n + \lambda_n x) = F(m_n) + \lambda_n (F_x) = \lambda_n F(x)$$

we have that $\lambda_n F(x) \to F(y)$. Since $F(x) \neq 0$, this implies that $\lambda_n \to F(x)^{-1} F(y)$. It follows that $\lambda_n x \to F(x)^{-1} F(y) x$. Since for each $n \in \mathbb{N}$, $m_n = y_n - \lambda_n x$, we know that $m_n \to y - F(x)^{-1} F(y) x$. Since $(m_n)_{n \in \mathbb{N}} \subset M$ and M is closed, we have that $y - F(x)^{-1} F(y) x \in M$ and therefore $y \in M + \mathbb{C}x$. Hence $M + \mathbb{C}x$ is closed.

(2) If $M = X$, then M is closed. Suppose that $M \neq X$. Let $(x_i)_{i=1}^n$ be a basis for M. Define $N_0 = \{0\}$ and for each $i = 1, 2, \cdots, n$, define $N_i = N_{i-1} + \mathbb{C}x_i$. Since N_0 is a proper closed subpace of X and $x_1 \in X \setminus N_0$, (1) implies that N_1 is closed. Proceed inductively to obtain that $M = N_n$ is closed.

$\square$

Exercise 4.5.24. Let X be an infinite-dimensional normed vector space.

(1) There exists a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that for each $m, n \in \mathbb{N}$, $\|x_n\| = 1$ and if $m \neq n$, then $\|x_m - x_n\| > \frac{1}{2}$.
(2) X is not locally compact.

Proof. (1) Define $N_0 = \{0\}$. Then N_0 is a closed proper subspace of X. Choose $x_1 \in X$ such that $\|x_1\| = 1$. Using the results of previous exercises, we proceed inductively. For each $n \geq 2$ we define $N_{n-1} = \text{span}(x_1, x_2, \cdots, x_{n-1})$. Then N_{n-1} is a closed proper subspace of X. Thus we may choose $x_n \in X$ such that $\|x_n\| = 1$ and $\|x_n + N_{n-1}\| > \frac{1}{2}$. Let $m, n \in \mathbb{N}$. Suppose that $m < n$. Then $x_m \in N_{n-1}$. Thus $\|x_n - x_m\| \geq \|x_n + N_{n-1}\| > \frac{1}{2}$

(2) Suppose that X is locally compact. Then $\overline{B(0,1)}$ is compact and therefore sequentially compact. Using $(x_n)_{n\in\mathbb{N}} \subset \overline{B(0,1)}$ defined in (1), we see that there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$, $x \in \overline{B(0,1)}$ such that $x_{n_k} \to x$. Then $(x_{n_k})_{k\in\mathbb{N}}$ is Cauchy. So there exists $N \in N$ such that for each $j, k \in \mathbb{N}$, if $j, k \geq N$, then $\|x_{n_j} - x_{n_k}\| < \frac{1}{2}$. Then $\|x_{n_N} - x_{n_{N+1}}\| < \frac{1}{2}$. This is a contradiction since by construction, $\|x_{n_N} - x_{n_{N+1}}\| > \frac{1}{2}$. Thus X is not locally compact.

□

Exercise 4.5.25. Let X, Y be normed vector spaces and $T \in L(X, Y)$.

(1) Define the **adjoint of** T, denoted $T^* : Y^* \to X^*$ by $T^*(f) = f \circ T$. Then $T^* \in L(Y^*, X^*)$.
(2) Applying the result from (1) twice, we have that $T^{**} \in L(X^{**}, Y^{**})$. We have that for each $x \in X$, $T^{**}(\hat{x}) = \widehat{T(x)}$.
(3) T^* is injective iff $T(X)$ is dense in Y.
(4) If $T^*(Y^*)$ is dense in X^*, then T is injective. The converse is true if X is reflexive.

Proof. (1) Let $f \in Y^*$. Then $\|T^*(f)\| = \|f \circ T\| \leq \|T\|\|f\|$. So $T^* \in L(Y^*, X^*)$ with $\|T^*\| \leq \|T\|$.

(2) Let $x \in X$. Let $f \in Y^*$. Then

$$\begin{aligned}
T^{**}(\hat{x})(f) &= \hat{x} \circ T^*(f) \\
&= \hat{x}(T^*(f)) \\
&= \hat{x}(f \circ T) \\
&= f \circ T(x) \\
&= f(T(x)) \\
&= \widehat{T(x)}(f)
\end{aligned}$$

Hence $T^{**}(\hat{x}) = \widehat{T(x)}$.

(3) Suppose that $T(X)$ is not dense in Y. Then $\overline{T(X)} \neq Y$. So $T(X)$ is a proper closed subspace of Y and there exists $y \in Y$ such that $y \notin \overline{T(X)}$. By a previous exercise, there exists $f \in Y^*$ such that $f(y) = \|y + \overline{T(X)}\| \neq 0$, $\|f\| = 1$ and $f|_{\overline{T(X)}} = 0$. Let $x \in X$. Then $T^*(f)(x) = f \circ T(x) = 0$. Hence $T^*(f) = 0 = T^*(0)$. Since $f \neq 0$, T^* is not injective.
Now suppose that $T(X)$ is dense in Y. Let $f, g \in Y^*$. Define $h \in Y^*$ by $h = f - g$ Suppose that $T * (f) = T^*(g)$ Then $T^*(h) = 0$. So for each $x \in X$, $h(T(x)) = 0$. Let $y \in Y$ and $\epsilon > 0$. By continuity, there exists $\delta > 0$ such that for each $y' \in Y$, if $\|y - y'\| < \delta$, then $\|h(y) - h(y')\| < \epsilon$. Since $T(X)$ is dense in Y, there exists $x \in X$ such that $\|y - T(x)\| < \delta$. Thus

$$\begin{aligned}\|h(y)\| &\leq \|h(y) - h(T(x))\| + \|h(T(x))\| \\ &= \|h(y) - h(T(x))\| \\ &< \epsilon\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, $\|h(y)\| = 0$. This implies that $h(y) = 0$ and therefore $f(y) = g(y)$. Since $y \in Y$ is arbitrary, $f = g$ and T^* is injective.

(4) For the sake of contradiction, suppose that $T^*(Y^*)$ is dense in X^* and T is not injective. Then there exist $x_1, x_2 \in X$ such that $x_1 \neq x_2$ and $T(x_1) = T(x_2)$. Define $x = x_1 - x_2$. Then $x \neq 0$ and $T(x) = 0$. A previous exercise implies that there exists $F \in X^*$ such that $F(x) = \|x\| \neq 0$ and $\|F\| = 1$. Let $\epsilon > 0$. Choose $g \in Y^*$ such that $\|F - T^*(g)\| < \epsilon$. Then

$$\begin{aligned}\|x\| &= |F(x)| \\ &\leq |F(x) - T^*(g)(x)| + |T^*(g)(x)| \\ &< \epsilon\|x\| + |g(T(x))| \\ &= \epsilon\|x\|\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have that $\|x\| = 0$ which is a contradiction. Hence if $T^*(Y^*)$ is dense in X^*, then T is injective.

Now, suppose that X is reflexive and T is injective. Let $\phi_1, \phi_2 \in X^{**}$. Suppose that $T^{**}(\phi_1) = T^{**}(\phi_2)$. Then $T^{**}(\phi_1 - \phi_2) = 0$. Since X is reflexive, there exist $x_1, x_2 \in X$ such that $\phi_1 = \hat{x_1}$ and $\phi_2 = \hat{x_2}$. Define $x = x_1 - x_2$. Then $T^{**}(\hat{x}) = 0$. So for each $f \in Y^*$,

$$\begin{aligned}T^{**}(\hat{x})(f) &= \hat{x} \circ T^*(f) \\ &= \hat{x}(T^*(f)) \\ &= \hat{x}(f \circ T) \\ &= f \circ T(x) \\ &= f(T(x)) \\ &= 0\end{aligned}$$

Suppose that $T(x) \neq 0$. Then a previous exercise implies that there exists $g \in Y^*$ such that $g(T(x)) = \|T(x)\| \neq 0$ and $\|g\| = 1$. This is a contradiction since $g(T(x)) = 0$.

So $T(x) = 0$. Since T is injective, this implies that $x = 0$. Hence $\hat{x} = 0$ and thus $\phi_1 = \phi_2$. Thus T^{**} is injective. By (3), we have that $T^*(Y^*)$ is dense in X^*.

$\square$

Exercise 4.5.26. Let X be a normed vector space. Then X is reflexive iff X^* is reflexive.

Proof. Suppose that X is reflexive. Let $\alpha \in X^{***}$. Define $f : X \to \mathbb{C}$ by $f(x) = \alpha(\hat{x})$. Clearly f is linear and a previous exercise tells us that for each $x \in X$,

$$\begin{aligned} |f(x)| &\leq \|\alpha\|\|\hat{x}\| \\ &= \|\alpha\|\|x\| \end{aligned}$$

So $f \in X^*$. Let $\phi \in X^{**}$. Since X is reflexive, there exists $x \in X$ such that $\phi = \hat{x}$. Then

$$\begin{aligned} \alpha(\phi) &= \alpha(\hat{x}) \\ &= f(x) \\ &= \hat{x}(f) \\ &= \hat{f}(\hat{x}) \\ &= \hat{f}(\phi) \end{aligned}$$

Hence $\alpha = \hat{f}$. Thus the map $X^* \to X^{***}$ given by $f \mapsto \hat{f}$ is surjective and so X^* is reflexive.

Conversely, suppose that X^* is reflexive. Since $\phi : X \to X^{**}$ given by $\phi(x) = \hat{x}$ is an isometry, $\widehat{X} \subset X^{**}$ is closed. For the sake of contradiction, suppose that $\widehat{X} \neq X^{**}$. Then there exists $\alpha \in X^{**}$ such that $\alpha \notin \widehat{X}$. Thus there exists $F \in X^{***}$ such that $\|F\| = 1$, $F(\alpha) = \|\alpha + \widehat{X}\| \neq 0$ and $F|_{\widehat{X}} = 0$. Since X^* is reflexive, there exists $f \in X^*$ such that $F = \hat{f}$. A previous exercise tells us that $\|f\| = \|\hat{f}\| = \|F\| = 1$. Since for each $x \in X$, $f(x) = \hat{x}(f) = \hat{f}(\hat{x}) = F(\hat{x}) = 0$, we have that $f = 0$. Thus $\|f\| = 0$, a contradiction. So $\widehat{X} = X^{**}$ and X is reflexive.

$\square$

4.6. **The Baire Category and Closed Graph Theorems.**

Theorem 4.6.1. *Let X, Y be Banach spaces and $T \in L(X, Y)$. If T is surjective, then T is open.*

Corollary 4.6.2. *Let X, Y be Banach spaces and $T \in L(X, Y)$. If T is a bijection, then $T^{-1} \in L(X, Y)$.*

Definition 4.6.1. Let X, Y be sets and $f : X \to Y$. We define the **graph of f**, $\Gamma(f)$, by $\Gamma(f) = \{(x, y) \in X \times Y : f(x) = y\}$.

Theorem 4.6.3. *Let X, Y be Banach spaces and $T : X \to Y$ a linear map. If $\Gamma(T)$ is closed, then $T \in L(X, Y)$.*

Note 4.6.1. We recall that $\Gamma(T)$ is closed iff for each $(x_n)_{n\in\mathbb{N}} \subset X$, $x \in X$ and $y \in Y$, $x_n \to x$ and $T(x_n) \to y$ implies that $T(x) = y$.

Theorem 4.6.4. *Let X, Y be Banach spaces and $S \subset L(X, Y)$. If for each $x \in X$,*

$$\sup_{T\in S} \|Tx\| < \infty$$

then

$$\sup_{T\in S} \|T\| < \infty$$

Exercise 4.6.2. Let μ be counting measure on $(N, \mathcal{P}(\mathbb{N}))$. Define $h : \mathbb{N} \to \mathbb{N}$ and ν on $(N, \mathcal{P}(\mathbb{N}))$ by $h(n) = n$ and $d\nu = hd\mu$. Define $X = L^1(\nu)$ and $Y = L^1(\mu)$. Equip both X and Y with the L^1 norm with respect to μ.

(1) We have that X is a proper subspace of Y and therefore X is not complete.
(2) Define $T : X \to Y$ by $Tf(n) = nf(n)$. Then T is linear, $\Gamma(T)$ is closed, and T is unbounded.
(3) Define $S : Y \to X$ by $Sg(n) = \frac{1}{n}g(n)$. Then $S \in L(Y, X)$, S is surjective and S is not open.

Proof.

(1) Note that for each $f : \mathbb{N} \to \mathbb{C}$,

$$\begin{aligned} \|f\|_{\mu,1} &= \sum_{n=1}^{\infty} |f(n)| \\ &\leq \sum_{n=1}^{\infty} n|f(n)| \\ &= \|f\|_{\nu,1} \end{aligned}$$

Hence X is a subspace of Y. Define $f : \mathbb{N} \to \mathbb{C}$ by $f(n) = \frac{1}{n^2}$. Then

$$\|f\|_{\mu,1} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

So $f \in Y$. However

$$\|f\|_{\nu,1} = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

So $f \notin X$. Thus X is a proper subspace of Y. Let $g \in Y$ and $\epsilon > 0$. Since the simple functions are dense in $L^1(\mu)$, there exists $\phi \in L^1(\mu)$ such that ϕ is simple and $\|g - \phi\|_{\mu,1} < \epsilon$. Then there exist $(c_i)_{i=1}^k \subset \mathbb{C}$ and $(E_i)_{i=1}^k \subset \mathcal{P}(\mathbb{N})$ such that for each $i, j \in \{1, 2, \cdots, k\}$, E_i is finite, $i \neq j$ implies that $E_i \cap E_j = \varnothing$ and

$$\phi = \sum_{i=1}^{k} c_i \chi_{E_i}$$

Define $c = \max\{|c_i| : i = 1, 2, \cdots k\}$ and $m = \max \left[\bigcup_{i=1}^k E_i \right]$. Then

$$\begin{aligned} \|\phi\|_{\nu,1} &= \sum_{n=1}^{m} n|\phi(n)| \\ &\leq \sum_{n=1}^{m} mc \\ &= cm^2 \\ &< \infty \end{aligned}$$

Hence $\phi \in X$ and X is dense in Y. Since X is a dense, proper subspace, it is not closed. Since Y is complete and $X \subset Y$ is not closed, we have that X is not complete.

(2) Clearly T is linear. Let $(f_j)_{j \in \mathbb{N}} \subset X$, $f \in X$ and $g \in Y$. Suppose that $f_j \xrightarrow{L^1(\mu)} f$ and $Tf_j \xrightarrow{L^1(\mu)} g$.

Note that for each $j \in \mathbb{N}$ and $n \in \mathbb{N}$,

$$|f_j(n) - f(n)| \leq \sum_{n=1}^{\infty} |f_j(n) - f(n)| = \|f_j - f\|_{\mu,1}$$

and

$$|nf_j(n) - g(n)| \leq \sum_{n=1}^{\infty} |nf_j(n) - g(n)| = \|Tf_j - g\|_{\mu,1}$$

Thus for each $n \in \mathbb{N}$, $f_j(n) \xrightarrow{j} f(n)$ and $nf_j(n) \xrightarrow{j} g(n)$. This implies that for each $n \in \mathbb{N}$, $nf(n) = g(n)$. Thus $Tf = g$ which implies that $\Gamma(T)$ is closed. Suppose, for the sake of contradiction, that T is bounded. Then there exists $C \geq 0$ such that for each $f \in X$, $\|Tf\|_{\mu,1} \leq C\|f\|_{\mu,1}$. Choose $n \in \mathbb{N}$ such that $n > C$. Define $f : \mathbb{N} \to \mathbb{C}$ by $f = \chi_{\{n\}}$. As established above, $S^+ \subset L^1(\mu)$. Then $\|f\|_{\mu,1} = 1$ and

$$\begin{aligned} \|Tf\|_{\mu,1} &= n \\ &> C \\ &= C\|f\|_{\mu,1} \end{aligned}$$

which is a contradiction. So T is unbounded.

(3) Clearly S is linear. Let $g \in Y$. Then

$$\begin{aligned}\|Sg\|_{\mu,1} &= \sum_{n=1}^{\infty} \frac{1}{n}|g(n)| \\ &\leq \sum_{n=1}^{\infty} |g(n)| \\ &= \|g\|_{\mu,1}\end{aligned}$$

So S is bounded and $\|S\| \leq 1$. Thus $S \in L(Y, X)$. Let $f \in X$. Define $g : \mathbb{N} \to \mathbb{C}$ by $g(n) = nf(n)$. By defnition, $g \in Y$ and we have that

$$\begin{aligned}Sg(n) &= \frac{1}{n}g(n) \\ &= f(n)\end{aligned}$$

Hence $Sg = f$ and thus S is surjective. Let $g \in Y$. Suppsose that $Sg = 0$. Then

$$\sum_{n=1}^{\infty} \frac{1}{n}|g(n)| = \|Sg\| = 0$$

Thus for each $n \in \mathbb{N}$, $g(n) = 0$. Hence $\ker S = \{0\}$ and S is injective. Note that for each $A \subset Y$, $S(A) = T^{-1}(A)$. If S is open, then T is continuous which as shown above is a contradiction. So g is not open.

$\square$

Exercise 4.6.3. Let $X = C^1([0,1])$ and $Y = C([0,1])$. Equip both X and Y with the uniform norm.

(1) Then X is not complete
(2) Define $T : X \to Y$ by $Tf = f'$. Then $\Gamma(T)$ is closed and T is not bounded.

Proof. (1) Recall that for each $a, b \geq 0$ and $p \in \mathbb{N}$,

$$(a^{\frac{1}{p}} + b^{\frac{1}{p}})^p = \sum_{n=0}^{p} \binom{p}{n} a^{\frac{n}{p}} b^{\frac{p-n}{p}} \geq a + b$$

Thus $(a+b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$.

For each $n \in \mathbb{N}$, define $f_n : [0,1] \to \mathbb{C}$ by $f_n(x) = \sqrt{(x - \frac{1}{2})^2 + \frac{1}{n^2}}$. Then $(f_n)_{n \in \mathbb{N}} \subset X$. Define $f : [0,1] \to \mathbb{C}$ by $f(x) = |x - \frac{1}{2}|$. Then $f \in Y \cap X^c$. Note that for each $n \in \mathbb{N}$, $f \leq f_n$. Our observation above implies that for each $x \in X$,

$$\begin{aligned}f_n(x) &= \left[(x - \frac{1}{2})^2 + \frac{1}{n^2}\right]^{\frac{1}{2}} \\ &\leq |x - \frac{1}{2}| + \frac{1}{n}\end{aligned}$$

Thus $0 \leq f_n - f \leq \frac{1}{n}$. This implies that $f_n \xrightarrow{u} f$. Since $f \notin X$, X is not complete.

(2) Let $(f_n)_{n\in\mathbb{N}} \subset X$, $f \in X$ and $g \in Y$. Suppose that $f_n \xrightarrow{u} f$ and $Tf_n \xrightarrow{u} g$. Let $x \in [0,1]$. Then $f_n(x) \to f(x)$ and $f_n(0) \to f(0)$ and $f_n' \xrightarrow{u} g$. Applying the DCT to this sequence of integrable functions that converges uniformly to an integrable function on a finite measure space (a previous exercise) we have that

$$\begin{aligned} f_n(x) - f_n(0) &= \int_{[0,x]} f_n' dm \\ &\to \int_{[0,x]} g dm \end{aligned}$$

Since $f_n(x) - f_n(0) \to f(x) - f(0)$, we know that

$$f(x) - f(0) = \int_{[0,x]} g dm$$

. Thus $Tf = g$ and $\Gamma(T)$ is closed.
Suppose for the sake of contradiction that T is bounded. Then there exists $C \geq 0$ such that for each $f \in X$, $\|Tf\| \leq C\|f\|$. Choose $n \in \mathbb{N}$ such that $n > C$. Define $f \in X$ by $f(x) = x^n$. Then $\|f\| = 1$ and

$$\begin{aligned} \|Tf\| &= \|f'\| \\ &= n \\ &> C \\ &= C\|f\| \end{aligned}$$

which is a contradiction. So T is not bounded.

□

Exercise 4.6.4. Let X, Y be Banach spaces and $T \in L(X,Y)$. Then $X/\ker T \cong T(X)$ iff $T(X)$ is closed.

Proof. Since X is a banach space and T is continuous, we have that $\ker T$ is closed and $X/\ker T$ is a Banach space. Suppose that $X/\ker T \cong T(X)$. Then $T(X)$ is complete. Since Y is complete, this implies that $T(X)$ is closed.
Conversely Suppose that $T(X)$ is closed. Then $T(X)$ is complete. Define $S : X/\ker T \to T(X)$ by $S(x+\ker T) = T(x)$. A previous exercise tells us that the map $S : X/\ker T \to T(X)$ defined by $S(x + \ker T) = T(x)$ is a bounded linear bijection. Since $T(X)$ is complete and S is surjective, S^{-1} is bounded and thus S is an isomorphism. □

Exercise 4.6.5. Let X be a separable Banach space. Define $B_X = \{x \in X : \|x\| < 1\}$. Let $(x_n)_{n\in\mathbb{N}} \subset B_X$ a dense subset of the unit ball and μ the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Define $T : L^1(\mu) \to X$ by

$$Tf = \sum_{n=1}^{\infty} f(n) x_n$$

Then

(1) T is well defined and $T \in L(L^1(\mu), X)$
(2) T is surjective
(3) There exists a closed subspace $K \subset L^1(\mu)$ such that $L^1(\mu)/K \cong X$

Proof. (1) Let $f \in L^1(\mu)$. Since X is complete and

$$\begin{aligned}\sum_{n=1}^{\infty} \|f(n)x_n\| &= \sum_{n=1}^{\infty} |f(n)|\|x_n\| \\ &\leq \sum_{n=1}^{\infty} |f(n)| \\ &< \infty\end{aligned}$$

we have that $\sum_{n=1}^{\infty} f(n)x_n$ converges and thus $Tf \in X$. Hence T is well defined.

Clearly T is linear. Let $f \in L^1(\mu)$. Then

$$\begin{aligned}\|Tf\| &= \|\sum_{n=1}^{\infty} f(n)x_n\| \\ &\leq \sum_{n=1}^{\infty} \|f(n)x_n\| \\ &\leq \sum_{n=1}^{\infty} |f(n)| \\ &= \|f\|_1\end{aligned}$$

So T is bounded with $\|T\| \leq 1$.

(2) Let $x \in X$. Suppose that $\|x\| < 1$. Then $x \in B_X$. So there exists $n_1 \in \mathbb{N}$ such that $\|x - x_{n_1}\| < \frac{1}{2}$. Then $2(x - x_{n_1}) \in B_X$. Since for each $j \in \mathbb{N}$, $B_X \setminus (x_n)_{n=1}^{j}$ is dense in B_X, there exists $n_2 \in \mathbb{N}$ such that $x_{n_2} \notin (x_n)_{n=1}^{n_1}$ and $\|2(x - x_{n_1}) - x_{n_2}\| < \frac{1}{2}$ which implies that $\|x - (x_{n_1} - \frac{1}{2}x_{n_2})\| < \frac{1}{4}$.

Proceed inductively to obtain a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that for each $k \geq 2$, $x_{n_k} \notin (x_n)_{n=1}^{n_{k-1}}$ and $\|x - \sum_{j=1}^{k} 2^{1-j}x_{n_j}\| < \frac{1}{2^k}$. Then $x = \sum_{k=1}^{\infty} 2^{1-k}x_{n_k}$.

Define $f : \mathbb{N} \to \mathbb{C}$ by $f = \sum_{k=1}^{\infty} 2^{1-k}\chi_{\{n_k\}}$. Then $\|f\|_1 = \sum_{k=1}^{\infty} 2^{1-k} < \infty$, so $f \in L^1(\mu)$ and $Tf = \sum_{k=1}^{\infty} 2^{1-k}x_{n_k} = x$. Now, suppose that $\|x\| \geq 1$, then $\frac{1}{2\|x\|}x \in B_X$. The above argument shows that there exists $f \in L^1(\mu)$ such that $Tf = \frac{1}{2\|x\|}x$. Then $2\|x\|f \in L^1(\mu)$ and $T(2\|x\|f) = 2\|x\|Tf = x$.
So for each $x \in X$, there exists $f \in L^1(\mu)$ such that $Tf = x$ and thus T is surjective.

(3) Since X is a Banach space and T is surjective, the previous exercise implies that $L^1(\mu)/\ker T \cong X$.

□

Exercise 4.6.6. Let X, Y be Banach spaces and $T : X \to Y$ a linear map. If for each $f \in Y^*$, $f \circ T \in X^*$, then $T \in L(X, Y)$.

Proof. Suppose that for each $f \in Y^*$, $f \circ T \in X^*$. Let $x \in X$, □

4.7. **Banach Algebras.**

Definition 4.7.1. Let X be a Banach space and an associative algebra. Then X is said to be a **Banach algebra** if for each $S, T \in X$, $\|ST\| \leq \|S\|\|T\|$.

Definition 4.7.2. Let X be a Banach algebra and $I \in X$. Then I is said to be an **identity** if for each $T \in X$, $IT = TI = T$.

Definition 4.7.3. Let X be a Banach algebra. and $I \in X$. Then I is said to be an **identity** if $I \neq 0$ and for each $T \in X$, $IT = TI = T$.

Definition 4.7.4. Let X be a Banach algebra. Then X is said to be **unital** if there exists $I \in X$ such that I is an identity.

Exercise 4.7.5. Let X be a unital Banach algebra. Then there exists a unique $I \in X$ such that I is an identity.

Proof. Clear. □

Note 4.7.1. We denote the unique identity element by I.

Definition 4.7.6. Let X be a unital Banach algebra and $T, S \in X$. Then S is said to be an **inverse** of T if $TS = ST = I$.

Definition 4.7.7. Let X be a unital Banach algebra and $T \in X$. Then T is said to be **invertible** if there exists $S \in X$ such that S is an inverse of T.

Exercise 4.7.8. Let X be a unital Banach algebra and $T \in X$. If T is invertible, then there exists a unique $S \in X$ such that S is an inverse of T.

Proof. Clear. □

Note 4.7.2. We denote the unique inverse of T by T^{-1}.

Exercise 4.7.9. Fundamental Example:
Let X be a Banach space. Then $GL(X)$ is a unital Banach algebra.

Proof. Clear. □

Definition 4.7.10. Let X be a unital Banach algebra. We define $GL(X) = \{T \in X : T \text{ is invertible}\}$.

Exercise 4.7.11. Let X be a unital Banach algebra. Then $GL(X)$ is a group.

Proof. Clear. □

Exercise 4.7.12. Let X be a unital Banach algebra. Then $1 \leq \|I\|$.

Proof. Since $I \neq 0$, $\|I\| \neq 0$. By definition,

$$\|I\| = \|II\| \leq \|I\|\|I\|$$

Hence $1 \leq \|I\|$. □

Exercise 4.7.13. Let X be a Banach algebra. Then mulitplication is continuous.

Proof. Let $(S_1, T_1) \in X \times X$ and $\epsilon > 0$. Choose $\delta = \min\{\frac{\epsilon}{2(\|S_1\|+\|T_1\|+1)}, \frac{\sqrt{\epsilon}}{\sqrt{2}}\}$. Let $(S_2, T_2) \in X \times X$. Suppose that

$$\|(S_1, T_1) - (S_2, T_2)\| = \max\{\|S_2 - S_2\|, \|T_1 - T_2\|\} < \delta$$

Then

$$\begin{aligned}
\|S_1T_1 - S_2T_2\| &= \|S_1T_1 - S_2T_1 + S_2T_1 - S_2T_2\| \\
&\leq \|S_1 - S_2\|\|T_1\| + \|S_2\|\|T_1 - T_2\| \\
&\leq \|S_1 - S_2\|\|T_1\| + \big(\|S_1 - S_2\| + \|S_1\|\big)\|T_1 - T_2\| \\
&\leq \delta\|T_1\| + (\delta + \|S_1\|)\delta \\
&= \delta(\|S_1\| + \|T_1\|) + \delta^2 \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

□

Exercise 4.7.14. Let X be a unital Banach algebra. Then

(1) For each $T \in X$, if $\|I - T\| < 1$, then $T \in GL(X)$ and

$$T^{-1} = \sum_{n=0}^{\infty}(I - T)^n$$

(2) For each $S, T \in X$, if $S \in GL(X)$ and $\|S - T\| < \|S^{-1}\|^{-1}$, then $T \in GL(X)$.
(3) $GL(X)$ is open.

Proof.

(1) Let $T \in X$. Suppose that $\|I - T\| < 1$. Then

$$\sum_{n=0}^{\infty} \|(I - T)^n\| \leq \sum_{n=0}^{\infty} \|I - T\|^n < \infty$$

Since X is a complete, $\sum\limits_{n=0}^{\infty}(I - T)^n$ converges in X.

Define $(S_k)_{k=0}^{\infty} \subset X$ and $S \in X$ by $S_k = \sum\limits_{n=0}^{k}(I - T)^n$ and $S = \sum\limits_{n=0}^{\infty}(I - T)^n$. Then for each $k \in \mathbb{N}$,

$$\begin{aligned}
S_kT &= S_k - S_k(I - T) \\
&= (I - T)^0 - (I - T)^{k+1} \\
&= I - (I - T)^{k+1}
\end{aligned}$$

and $\|S_kT - I\| \leq \|I - T\|^{k+1}$. Since multiplication on Banach algebras is continuous, we have that

$$ST = (\lim_{k\to\infty} S_k)T = \lim_{k\to\infty} S_kT = I$$

Similarly $TS = I$. Thus $T \in GL(X)$ and $T^{-1} = S \in X$.

(2) Let $S, T \in X$. Suppose that $S \in GL(X)$ and $\|S - T\| < \|S^{-1}\|^{-1}$. Then

$$\begin{aligned}\|I - S^{-1}T\| &= \|S^{-1}(S - T)\| \\ &\leq \|S^{-1}\|\|S - T\| \\ &< 1\end{aligned}$$

So $S^{-1}T \in GL(X)$. Thus $T = S(S^{-1}T) \in GL(X)$.

(3) Let $T \in GL(X)$. Choose $\delta = \|T^{-1}\|^{-1}$. By (2), $B(T, \delta) \subset GL(X)$.

$\square$

5. Hilbert Spaces

5.1. Introduction.

Definition 5.1.1. Let H be a vector space and $\langle \cdot, \cdot \rangle : H \to \mathbb{C}$. Then $\langle \cdot, \cdot \rangle$ is said to be an **inner product** on H if for each $x, y, z \in H$and $c \in \mathbb{C}$

(1) $\langle x, y + cz \rangle = \langle x, y \rangle + c\langle x, z \rangle$
(2) $\langle x, y \rangle = \langle y, x \rangle^*$
(3) $\langle x, x \rangle \geq 0$
(4) if $\langle x, x \rangle = 0$, then $x = 0$.

Note 5.1.1. In mathematics, inner products are conventionally defined to be linear in the first argument. However, in my opinion, the convention in physics of defining inner products to be linear in the second argument makes more sense.

Exercise 5.1.2. Let H be an inner product space, $(x_j)_{j=1}^n$, $(y_j)_{j=1}^n \subset H$ and $(\alpha_j)_{j=1}^n$, $(\beta_j)_{j=1}^n \subset \mathbb{C}$. Then

$$\left\langle \sum_{i=1}^n \alpha_i x_i, \sum_{j=1}^n \beta_j y_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_i^* \beta_j \langle x_i, y_j \rangle$$

Proof. Clear. □

Definition 5.1.3. Let H be an inner product space. Define the **induced norm**, denoted $\| \cdot \| : H \to \mathbb{C}$, by

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}}$$

Exercise 5.1.4. Cauchy-Schwarz Inequality
Let H be an inner product space. Then for each $x, y \in H$, $|\langle x, y \rangle| \leq \|x\|\|y\|$ and $|\langle x, y \rangle| = \|x\|\|y\|$ iff $x \in \operatorname{span}(y)$.
Hint: For $x, y \in H$, put $z = \operatorname{sgn}\langle x, y \rangle^* y$ and Consider $f : \mathbb{R} \to [0, \infty)$ defined by $f(t) = \|x - tz\|^2$

Proof. Let $x, y \in H$. If $y = 0$, then the claim holds trivially. Suppose that $y \neq 0$. Put $z = \operatorname{sgn}\langle x, y \rangle^* y$. So $\langle x, z \rangle = |\langle x, y \rangle|$ and $\|z\| = \|y\|$. Define $f : \mathbb{R} \to [0, \infty)$ by

$$f(t) = \|x - tz\|^2$$

. Then for each $t \in \mathbb{R}$,

$$\begin{aligned} 0 &\leq f(t) \\ &= \|x - tz\|^2 \\ &= \|x\|^2 + |t|^2\|z\|^2 - 2\operatorname{Re}(t\langle x, z \rangle) \\ &= \|x\|^2 + t^2\|y\|^2 - 2t|\langle x, y \rangle| \end{aligned}$$

Thus f is a quadratic with a minimum at $t_0 = \frac{|\langle x,y\rangle|}{\|y\|^2}$. Hence

$$\begin{aligned} 0 &\leq f(t_0) \\ &= \|x\|^2 + \frac{|\langle x,y\rangle|}{\|y\|^2} - 2\frac{|\langle x,y\rangle|}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle x,y\rangle|}{\|y\|^2} \end{aligned}$$

Which implies that

$$|\langle x,y\rangle|^2 \leq \|x\|^2\|y\|^2$$

and hence the claim holds. Clearly if $x \in \text{span}(y)$, then equality holds. Conversely, if equality holds, then $x - z = 0$ which implies that $x \in spn(y)$. □

Exercise 5.1.5. Let H be an inner product space. Then the induced norm, $\|\cdot\| : H \to \mathbb{C}$, is a norm.

Proof. Let $x, y \in H$ and $c \in \mathbb{C}$. Then

(1) By definition, if $\|x\| = 0$, then $\langle x,x\rangle = 0$, which implies that $x = 0$.
(2) Note that

$$\begin{aligned} \|cx\|^2 &= \langle cx, cx\rangle \\ &= c * c\langle x,x\rangle \\ &= |c|^2\|x\|^2 \end{aligned}$$

So $\|cx\| = |c|\|x\|$
(3) The Cauchy-Schwarz inequality implies that

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\,\text{Re}(\langle x,y\rangle) \\ &\leq \|x\|^2 + \|y\|^2 + 2|\langle x,y\rangle| \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

Hence $\|x+y\| \leq \|x\| + \|y\|$.

□

Definition 5.1.6. Let H be an inner product space, $x, y \in H$ and $S \subset H$. Then

(1) x and y are said to be **orthogonal** if $\langle x,y\rangle = 0$.
(2) S is said to be **orthogonal** if for each $x, y \in S$, x, y are orthogonal.

Exercise 5.1.7. (Pythagorean theorem):
Let H be an inner product space and $(x_j)_{j=1}^n \subset H$ an orthogonal set. Then

$$\left\| \sum_{j=1}^n x_j \right\|^2 = \sum_{j=1}^n \|x_j\|^2$$

Proof. We have that

$$\begin{aligned}
\Big\| \sum_{j=1}^n x_j \Big\|^2 &= \Big\langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j \Big\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle x_j, x_j \rangle \\
&= \sum_{j=1}^n \langle x_j, x_j \rangle \\
&= \sum_{j=1}^n \|x_j\|^2
\end{aligned}$$

$\square$

Exercise 5.1.8. Let H be an inner product space and $S \subset H$. Suppose that $0 \notin S$. If S is orthogonal, then S is linearly independent.

Proof. Let $x_1, \cdots, x_n \in S$ and $c_1, \cdots, c_n \in \mathbb{C}$. Suppose that $\sum_{j=1}^n c_j x_j = 0$. Since $(c_j x_j)_{j=1}^n$ is orthogonal, the Pythagorean theorem implies that

$$\begin{aligned}
0 &= \Big\| \sum_{i=1}^n c_i x_i \Big\| \\
&= \sum_{j=1}^n |c_j|^2 \|x_j\|
\end{aligned}$$

So for each $j \in \{1, \cdots, n\}$, $c_j = 0$ and S is linearly independent. $\square$

Definition 5.1.9. Let H be an inner product space and $S \subset H$. Then S is said to be **orthonormal** if S is orthogonal and for each $x \in S$, $\|x\| = 1$.

Exercise 5.1.10. Bessel's Inequality:
Let H be an inner product space and $S \subset H$. If S is orthonormal, then for each $x \in H$,

$$\sum_{u \in S} |\langle u, x \rangle|^2 \leq \|x\|$$

and in particular, $\{u \in S : \langle u, x \rangle \neq 0\}$ is countable.

Proof. Suppose that S is orthonormal. Let $x \in H$ and $F \subset S$ finite. Then the Pythagorean theorem implies that

$$\begin{aligned} 0 &\leq \left\| x - \sum_{u \in F} \langle u, x \rangle u \right\|^2 \\ &= \|x\|^2 + \left\| \sum_{u \in F} \langle u, x \rangle u \right\|^2 - 2 \operatorname{Re} \left\langle x, \sum_{u \in F} \langle u, x \rangle u \right\rangle \\ &= \|x\|^2 + \sum_{u \in F} |\langle u, x \rangle|^2 \|u\|^2 - 2 \sum_{u \in F} |\langle u, x \rangle|^2 \\ &= \|x\|^2 - \sum_{u \in F} |\langle u, x \rangle|^2 \end{aligned}$$

So

$$\sum_{u \in F} |\langle u, x \rangle|^2 \leq \|x\|$$

By definition of the sum,

$$\sum_{u \in S} |\langle u, x \rangle|^2 \leq \|x\|$$

Basic integration theory then tells us that $\{u \in S : \langle u, x \rangle \neq 0\}$ is countable. □

Definition 5.1.11. Let H be an inner product space. Then H is said to be a **Hilbert space** if H is a complete with respect to the induced norm on H.

Exercise 5.1.12. Let H be a Hilbert space and $S \subset H$. Suppose that S is orthonormal. Then the followong are equivalent:

(1) For each $x \in H$, if for each $u \in S$, $\langle u, x \rangle = 0$, then $x = 0$.
(2) For each $x \in H$, there exist $(u_j)_{j \in \mathbb{N}} \subset S$ such that $x = \sum\limits_{j \in \mathbb{N}} \langle u_j, x \rangle u_j$ and for each $u \notin (u_j)_{j \in \mathbb{N}}$, $\langle u, x \rangle = 0$.
(3) For each $x \in H$, $\|x\|^2 = \sum\limits_{u \in S} |\langle u, x \rangle|^2$.

Proof.

- (1) $\implies$ (2):
  Suppose that for each $x \in H$, if for each $u \in S$, $\langle u, x \rangle = 0$, then $x = 0$. Let $x \in H$. Put $S_* = \{u \in S : \langle u, x \rangle \neq 0\}$. The previous exercise implies that S_* is countable. Write $S_* = (u_j)_{j=1}^n$. The previous exercise tells us that $\sum\limits_{j \in \mathbb{N}} |\langle u_j, x \rangle|^2 \leq \|x\|^2$ and hence converges. Thus for $\epsilon > 0$, there exist $N \in \mathbb{N}$ such that for each $m, n \in \mathbb{N}$, $m, n \geq N$ implies that if $m < n$, then

$$\sum_{m+1}^{n} |\langle u_j, x \rangle|^2 < \epsilon$$

  Define $(y_n)_{n \in \mathbb{N}} \subset H$ by

$$y_n = \sum_{j=1}^{n} \langle u_j, x \rangle u_j$$

Then for each $m, n \in \mathbb{N}$, $m, n \geq N$ implies that if $m < n$, then

$$\begin{aligned}\|y_n - y_m\|^2 &= \left\| \sum_1^n \langle u_j, x\rangle u_j - \sum_1^m \langle u_j, x\rangle u_j \right\|^2 \\ &= \left\| \sum_{m+1}^n \langle u_j, x\rangle u_j \right\|^2 \\ &= \sum_{m+1}^n |\langle u_j, x\rangle|^2 \\ &< \epsilon\end{aligned}$$

So $(y_n)_{n\in\mathbb{N}}$ is Cauchy. Since H is complete, there exists $y \in H$ such that $y_n \to y$. By definition,

$$y = \sum_{j\in\mathbb{N}} \langle u_j, x\rangle u_j$$

Continuity of $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ implies that
(1) for each $u \in S \setminus S_*$,

$$\begin{aligned}\langle u, x - y\rangle &= \langle u, x\rangle - \langle u, y\rangle \\ &= \langle u, x\rangle - \lim_{n\to\infty} \langle u, y_n\rangle \\ &= \langle u, x\rangle - \lim_{n\to\infty} \sum_{j=1}^n \langle u_j, x\rangle \langle u, u_j\rangle \\ &= 0 - 0 \\ &= 0\end{aligned}$$

(2) for each $k \in \mathbb{N}$,

$$\begin{aligned}\langle u_k, x - y\rangle &= \langle u_k, x\rangle - \langle u_k, y\rangle \\ &= \langle u_k, x\rangle - \lim_{n\to\infty} \langle u_k, y_n\rangle \\ &= \langle u_k, x\rangle - \lim_{n\to\infty} \sum_{j=1}^n \langle u_j, x\rangle \langle u_k, u_j\rangle \\ &= \langle u_k, x\rangle - \langle u_k, x\rangle \\ &= 0\end{aligned}$$

So for each $u \in S$, $\langle u, x - y\rangle = 0$. By assumption, $x - y = 0$ and hence

$$x = \sum_{j\in\mathbb{N}} \langle u_j, x\rangle u_j$$

- (2) $\implies$ (3):
  Suppose that for each $x \in H$, there exist $(u_j)_{j\in\mathbb{N}} \subset S$ such that $x = \sum\limits_{j\in\mathbb{N}} \langle u_j, x\rangle u_j$ and

for each $u \notin (u_j)_{j\in\mathbb{N}}$, $\langle u, x\rangle = 0$. Then continuity of $\|\cdot\| : H \to [0,\infty)$ implies that

$$\begin{aligned}\|x\|^2 &= \left\| \lim_{n\to\infty} \sum_{j=1}^{n} \langle u_j, x\rangle u_j \right\|^2 \\ &= \lim_{n\to\infty} \left\| \sum_{j=1}^{n} \langle u_j, x\rangle u_j \right\|^2 \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} |\langle u_j, x\rangle|^2 \\ &= \sum_{j\in\mathbb{N}} |\langle u_j, x\rangle|^2 \\ &= \sum_{u\in S} |\langle u, x\rangle|^2\end{aligned}$$

- (3) $\implies$ (4):
  Suppose that for each $x \in H$, $\|x\|^2 = \sum\limits_{u\in S} |\langle u, x\rangle|^2$. Let $x \in H$. Suppose that for each $u \in S$, $\langle u, x\rangle = 0$. Then

$$\begin{aligned}\|x\|^2 &= \sum_{u\in S} |\langle u, x\rangle|^2 \\ &= 0\end{aligned}$$

  So $x = 0$

□

Definition 5.1.13. Let H be a Hilbert space and $S \subset H$. Then S is said to be an **orthonormal basis of** H if

(1) S is orthonormal
(2) for each $x \in H$, if for each $u \in S$, $\langle u, x\rangle = 0$, then $x = 0$

5.2. **Operators and Functionals.**

Definition 5.2.1. (Adjoint of an Operator):
Let H be a Hilbert space and $A, B \in L(H)$. Then B is said to be the **adjoint** of A if for each x_1, $x_2 \in H$,

$$\langle x_1, Ax_2 \rangle = \langle Bx_1, x_2 \rangle$$

In this case, we write

$$B = A^*$$

Note 5.2.1. In physics, the adjoint of A is typically denoted by $A^\dagger$.

Exercise 5.2.2. Let H be a Hilbert space, $A, B \in L(H)$ and $\lambda \in \mathbb{C}$, then

(1) $(A^*)^* = A$
(2) $(A + B)^* = A^* + B^*$
(3) $(AB)^* = B^*A^*$
(4) $(\lambda A)^* = \lambda^* A^*$
(5) A and B commute iff A^* and B^* commute.

Proof. Let x_1, $x_2 \in H$. Then

(1)

$$\begin{aligned} \langle Ax_1, x_2 \rangle &= \langle x_2, Ax_1 \rangle^* \\ &= \langle A^* x_2, x_1 \rangle^* \quad \text{(by definition)} \\ &= \langle x_1, A^* x_2 \rangle \end{aligned}$$

(2)

$$\begin{aligned} \langle x_1, (A + B)x_2 \rangle &= \langle x_1, Ax_2 \rangle + \langle x_1, Bx_2 \rangle \\ &= \langle A^* x_1, x_2 \rangle + \langle B^* x_1, x_2 \rangle \\ &= \langle (A^* + B^*)x_1, x_2 \rangle \end{aligned}$$

(3)

$$\begin{aligned} \langle x_1, ABx_2 \rangle &= \langle A^* x_1, Bx_2 \rangle \\ &= \langle B^* A^* x_1, x_2 \rangle \end{aligned}$$

(4)

$$\begin{aligned} \langle x_1, \lambda Ax_2 \rangle &= \lambda \langle x_1, Ax_2 \rangle \\ &= \lambda \langle A^* x_1, x_2 \rangle \\ &= \langle \lambda^* A^* x_1, x_2 \rangle \end{aligned}$$

(5) If A and B commute, then

$$\begin{aligned} A^* B^* &= (BA)^* \\ &= (AB)^* \\ &= B^* A^* \end{aligned}$$

Conversely, if A^* and B^* commute then

$$\begin{aligned} AB &= (B^*A^*)^* \\ &= (A^*B^*)^* \\ &= BA \end{aligned}$$

□

Definition 5.2.3. Let H be a Hilbert space and $Q \in L(H)$. Then Q is said to be **self-adjoint** if

$$Q = Q^*$$

Exercise 5.2.4. Let H be a Hilbert space and $Q \in L(H)$. If Q is a self-adjoint then

(1) the eigenvalues of Q are real.
(2) the eigenvectors of Q corresponding to distinct eigenvalues are orthogonal.

Proof. Suppose that Q is self-adjoint.

(1) Let λ be an eigenvalue of Q with corresponding eigenvector x. Then

$$\begin{aligned} \lambda\langle x, x\rangle &= \langle x, Qx\rangle \\ &= \langle Qx, x\rangle \\ &= \lambda^*\langle x, x\rangle \end{aligned}$$

Thus $\lambda = \lambda^*$ and is real

(2) Let λ_1 and λ_2 be eigenvalues of Q with corresponding eigenvectors x_1 and x_2. Suppose that $\lambda_1 \neq \lambda_2$. Then

$$\begin{aligned} \lambda_2\langle x_1, x_2\rangle &= \langle x_1, Qx_2\rangle \\ &= \langle Qx_1, x_2\rangle \\ &= \lambda_1\langle x_1, x_2\rangle \end{aligned}$$

So $(\lambda_2 - \lambda_1)\langle x_1, x_2\rangle = 0$. Which implies that $\langle x_1, x_2\rangle = 0$

□

Exercise 5.2.5. Let H be a Hilbert space, $A, B \in L(H)$ and $\lambda \in \mathbb{R}$. Suppose that A, B are self-adjoint. If A and B commute and then λAB is self-adjoint.

Proof.

$$\begin{aligned} (\lambda AB)^* &= \lambda^*(AB)^* \\ &= \lambda B^*A^* \\ &= \lambda BA \\ &= \lambda AB \end{aligned}$$

□

Definition 5.2.6. (Adjoint of a Vector):
Let H be a Hilbert space and $x \in H$. We define the **adjoint** of x, denoted $x^* \in H^*$, by $x^*y = \langle x, y\rangle$.

Note 5.2.2. In mathematics, where linearity of the inner product is in the first argument, x^* is typically referred to by $u_x \in H^*$ where $u_x(y) = \langle y, x\rangle$. In physics, where the inner product with linearity in the second argument, $x^*\phi$ is usually written in the so-called "bra-ket" notation as $\langle x|\phi\rangle$ which works smoothly since it aligns with the linearity of $u_x(\phi_1 + \lambda\phi_2)$ and the conjugate-linearity of $u_{x_1+\lambda x_2}(\phi)$. In this way, it generalizes the notation for $\langle x, y\rangle = x^T y$ for $\mathbb{R}^n$ to $\langle x, y\rangle = x^* y^*$ for $\mathbb{C}^n$.

Exercise 5.2.7. Let H be a Hilbert space, $x, y \in H$ and $\lambda \in \mathbb{C}$. Then

(1) $(x + y)^* = x^* + y^*$
(2) $(\lambda x)^* = \lambda^* x^*$

Proof. Clear. □

Definition 5.2.8. Let H be a Hilbert space, $x, y \in H$ and $A \in L(H)$. We define

(1) $x^*A \in H^*$ by $(x^*A)y = x^*(Ay)$
(2) $xy^* \in L(H)$ by $(xy^*)z = (y^*z)x$

Exercise 5.2.9. Let H be a Hilbert space, $A \in L(H)$ and $x \in H$. Then

$$(Ax)^* = x^*A^*$$

Proof. Let $y \in H$. Then

$$\begin{aligned}(Ax)^*y &= \langle Ax, y\rangle \\ &= \langle x, A^*y\rangle \\ &= x^*A^*y\end{aligned}$$

□

Definition 5.2.10. (Commutator):
Let H be a Hilbert space and $A, B \in L(H)$. The **commutator** of A and B, denoted $[A, B]$, is defined by

$$[A, B] = AB - BA$$

Exercise 5.2.11. Let H be a Hilbert space and $A, B, C \in L(H)$. Then

(1) $[AB, C] = A[B, C] + [A, C]B$
(2) $[A, BC] = B[A, C] + [A, B]C$

Proof.

(1)

$$\begin{aligned}[AB, C] &= ABC - CAB \\ &= ABC - ACB + ACB - CAB \\ &= A(BC - CB) + (AC - CA)B \\ &= A[B, C] + [A, C]B\end{aligned}$$

(2) Similar to (1).

□

6. Differentiation

6.1. The Gateaux Derivative.

Note 6.1.1. In this section, we assume all Banach spaces to be over $\mathbb{R}$.

Definition 6.1.1. Let X, Y be a Banach spaces, $A \subset X$ open, $f : A \to Y$, $x_0 \in A$ and $x \in X$. Then f is said to be

(1) **right-hand-differentiable at x_0 in the direction x** if the limit
$$\lim_{t \to 0^+} \frac{f(x_0 + tx) - f(x_0)}{t}$$
exists. If f is right-hand-differentiable at x_0 in the direction x, we define the **right-hand derivative** of f at x_0 in the direction x, denoted by $d^+f(x_0; x)$, to be the above limit.

(2) **left-hand-differentiable at x_0 in the direction x** if the limit
$$\lim_{t \to 0^-} \frac{f(x_0 + tx) - f(x_0)}{t}$$
exists. If f is right-hand-differentiable at x_0 in the direction x, we define the **left-hand derivative** of f at x_0 in the direction x, denoted by $d^-f(x_0; x)$, to be the above limit.

(3) **differentiable at x_0 in the direction x** if the limit
$$\lim_{t \to 0} \frac{f(x_0 + tx) - f(x_0)}{t}$$
exists. If f is differentiable at x_0 in the direction x, we define the **derivative** of f at x_0 in the direction x, denoted by $df(x_0; x)$, to be the above limit.

Exercise 6.1.2. Let X, Y be Banach spaces, $A \subset X$ open, $f : A \to \mathbb{R}$ and $x_0 \in A$. Then $df(x_0; 0) = 0$.

Proof. Clear. □

Definition 6.1.3. The Gateaux Derivative:
Let X, Y be Banach spaces, $A \subset X$ open, $f : A \to Y$ and $x_0 \in A$. Then f is said to be

(1) **right-hand Gateaux differentiable** at x_0 if for each $x \in X$, $d^+f(x_0; x)$ exits. We define the **right-hand Gateaux derivative** of f at x_0, denoted $d^+f(x_0) : X \to \mathbb{R}$, to be
$$d^+f(x_0)(x) = d^+f(x_0; x)$$

(2) **left-hand Gateaux differentiable** at x_0 if for each $x \in X$, $d^-f(x_0; x)$ exits. We define the **left-hand Gateaux derivative** of f at x_0, denoted $d^-f(x_0) : X \to \mathbb{R}$, to be
$$d^-f(x_0)(x) = d^-f(x_0; x)$$

(3) **Gateaux differentiable** at x_0 if for each $x \in X$, $df(x_0; x)$ exits. We define the **Gateaux derivative** of f at x_0, denoted $df(x_0) : X \to \mathbb{R}$, to be
$$df(x_0)(x) = df(x_0; x)$$

Definition 6.1.4. Let Y be a Banach space, $A \subset \mathbb{R}$ open and $f : A \to Y$. Then f is said to be **Gateaux differentiable** if for each $x \in A$, f is Gateaux differentiable at x. If f is Gateaux differentiable, we define $df : A \to Y^X$ by $x_0 \mapsto df(x_0)$.

Exercise 6.1.5. Let X, Y be Banach spaces, $A \subset X$ open, $f, g : A \to Y$, $\lambda \in \mathbb{R}$ and $x_0 \in A$. If f, g are Gateaux differentiable at x_0, then $f + \lambda g$ Gateaux differentiable at x_0 and $d[f + \lambda g](x_0) = df(x_0) + \lambda dg(x_0)$.

Proof. Similar to the case of the derivative from Calc I. □

Exercise 6.1.6. Let X, Y be Banach spaces, $A \subset X$ open, $f : A \to Y$ and $x_0 \in A$. Suppose that f is Gateaux differentiable at x_0. Then for each $\lambda \in \mathbb{R}$ and $x \in X$,

$$df(x_0)(\lambda x) = \lambda df(x_0)(x)$$

Proof. Let $\lambda \in \mathbb{R}$ and $x \in X$. Then

$$\begin{aligned} df(x_0)(\lambda x) &= \lim_{t \to 0} \frac{f(x_0 + t\lambda x) - f(x_0)}{t} \\ &= \lim_{t \to 0} \lambda \frac{f(x_0 + t\lambda x) - f(x_0)}{\lambda t} \\ &= \lambda \lim_{t \to 0} \frac{f(x_0 + t\lambda x) - f(x_0)}{\lambda t} \\ &= \lambda \lim_{t \to 0} \frac{f(x_0 + tx) - f(x_0)}{t} \\ &= \lambda df(x_0)(x) \end{aligned}$$

□

Exercise 6.1.7. Let X be a Banach space, $A \subset \mathbb{R}$ open, $f : A \to \mathbb{R}$ and $x_0 \in A$. Suppose that f is Gateaux differentiable at x_0. Then $df(x_0) \in L(\mathbb{R}, Y)$.

Proof. Let $x, y, lam \in \mathbb{R}$.

(1) The previous exercise implies

$$\begin{aligned} df(x_0)(x + \lambda y) &= df(x_0)((x + \lambda y)1) \\ &= (x + \lambda y) df(x_0)(1) \\ &= x df(x_0)(1) + \lambda y df(x_0)(1) \\ &= df(x_0)(x) + \lambda df(x_0)(y) \end{aligned}$$

So $df(x_0) : \mathbb{R} \to Y$ is linear.

(2) Since

$$\begin{aligned} \|df(x_0)(x)\| &= \|x df(x_0)(1)\| \\ &= |x| \|df(x_0)(1)\| \end{aligned}$$

We have that $df(x_0) : \mathbb{R} \to Y$ is bounded with $\|df(x_0)\| \leq \|df(x_0)(1)\|$.

□

Exercise 6.1.8. Let X be a Banach space, $A \subset X$ open, $f : A \to \mathbb{R}$ and $x_0 \in A$. If f is Gateaux differentiable at x_0 and f has a local extremum at x_0, then $df(x_0) = 0$.

Proof. Suppose that f is Gateaux differentiable at x_0 and f has a local minimum point at x_0. Then there exists $\delta > 0$ such that $B(x_0, \delta) \subset A$ and for each $y \in B(x_0, \delta)$, $f(x_0) \leq f(y)$. For the sake of contradiction, suppose that $df(x_0) \neq 0$. Then there exists $x \in X$ such that

$x \neq 0$ and $df(x_0)(x) \neq 0$.
First, suppose that $df(x_0)(x) < 0$. Choose $\epsilon = -df(x_0)(x) > 0$. Then there exists $t_0 > 0$ such that for each $t \in B^*(0, t_0)$, $x_0 + tx \in B(x_0, \delta)$ and

$$\left| \frac{f(x_0 + tx) - f(x_0)}{t} - df(x_0)(x) \right| < \epsilon$$

This implies that for each $t \in B^*(0, t_0)$,

$$\begin{aligned} \frac{f(x_0 + tx) - f(x_0)}{t} &< \epsilon + df(x_0)(x) \\ &= 0 \end{aligned}$$

and hence $f(x_0 + tx) < f(x_0)$, which is a contradiction.
Now, suppose that $df(x_0)(x) > 0$. Then

$$\begin{aligned} df(x_0)(-x) &= -df(x_0)(x) \\ &< 0 \end{aligned}$$

Similarly to above, this implies that there exists $t_0 > 0$ such that for each $t \in B^*(0, t_0)$, $x_0 - tx \in B(x_0, \delta)$ and $f(x_0 - tx) < f(x_0)$ which is a contradiction. So $df(x_0)(x) = 0$ and $df(x_0) = 0$.
If f has a local maximum at x_0, then $-f$ has a local minimum point at x_0. Then

$$\begin{aligned} df(x_0) &= -d[-f](x_0) \\ &= -0 \\ &= 0 \end{aligned}$$

$\square$

6.2. **The Frechet Derivative.**

Exercise 6.2.1. Let X, Y be a normed vector spaces and $\phi : X \to Y$ linear. If $\phi(h) = o(\|h\|)$ as $h \to 0$, then $\phi = 0$.

Proof. Let $h_0 \in X$. If $h_0 = 0$, then $\phi(h_0) = 0$. Suppose that $h_0 \neq 0$. Define $(h_n)_{n\in\mathbb{N}} \subset X$ by

$$h_n = \frac{h_0}{n}$$

Then $h_n \to 0$. By continuity of ϕ and our initial assumption we have that

$$\begin{aligned}\|h_0\|^{-1}\phi(h_0) &= \phi\left(\frac{h_0}{\|h_0\|}\right) \\ &= \phi\left(\frac{h_n}{\|h_n\|}\right) \\ &= \frac{\phi(h_n)}{\|h_n\|} \\ &\to 0\end{aligned}$$

which implies that $\|h_0\|^{-1}\phi(h_0) = 0$. So $\phi(h_0) = 0$ and hence $\phi = 0$. □

Exercise 6.2.2. Let X, Y be a normed vector spaces, $A \subset X$ open, $f : A \to Y$ and $x_0 \in A$. Suppose that there exists $\phi : X \to Y$ such that ϕ is linear and

$$f(x_0 + h) = f(x_0) + \phi(h) + o(\|h\|) \qquad \text{as } h \to 0$$

then ϕ is unique.

Proof. Suppose that there exists $\psi : X \to Y$ such that ψ is linear and such that

$$f(x_0 + h) = f(x_0) + \psi(h) + o(\|h\|) \qquad \text{as } h \to 0$$

Then $\phi(h) - \psi(h) = o(h)$. Since $\phi - \psi$ is linear, the previous exercise implies that $\phi = \psi$. □

Note 6.2.1. Recall that for Banach spaces X and Y, there isomorphic isometry

$$L(X, L(X, \cdots, L(X, Y)) \cdots) \to L^n(X, Y)$$

given by $\phi \mapsto \psi_\phi$ where

$$\psi_\phi(x_1, x_2, \cdots, x_n) = \phi(x_1)(x_2), \cdots, (x_n)$$

Definition 6.2.3. Frechet Derivative:
Let X, Y be a banach spaces, $A \subset X$ open, $f : A \to Y$ and $x_0 \in A$.

(1)
- Then f is said to be **Frechet differentiable at** x_0 if there exists $Df(x_0) \in L(X, Y)$ such that,
$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + o(\|h\|) \qquad \text{as } h \to 0$$
- If f is Frechet differentiable at x_0, we define the **Frechet derivative of** f **at** x_0 to be $Df(x_0)$.
- We say that f is **Frechet differentiable** if for each $x \in A$, f is Frechet differentiable at x.
- If f is Frechet differentiable, we define the **Frechet derivative of** f, denoted $Df : A \to L(X, Y)$, by $x \mapsto D^{(1)}f(x)$.

(2) Continuing inductively, we set $D^0 f = f$ and for $n \geq 2$,

- f is said to be **n-th order Frechet differentiable at** x_0 if f is $(n-1)$-th order Frechet differentiable and $D^{n-1}f$ is Frechet differentiable at x_0.
- If f is n-th order Frechet differentiable at x_0, we define $D^n f(x_0) \in L^n(X,Y)$ by

$$D^n f(x_0) = D[D^{n-1}f](x_0)$$

- We say that f is **n-th order Frechet differentiable** if f is $(n-1)$-th order Frechet differentiable and for each $x \in A$, $D^{n-1}f$ is Frechet differentiable at x.
- If f is n-th order Frechet differentiable, we define the **n-th order Frechet derivative of** f, denoted $D^n f : A \to L^n(X,Y)$ by $x \mapsto D^n f(x)$

(3) If f is n-th order differentiable, then f is said to be **continuously n-th order differentiable** if $D^n f$ is continuous. We define

$$C^n(A,Y) = \{f : A \to Y : f \text{ is continuously } n\text{-th order differentiable}\}$$

Exercise 6.2.4. Let X, Y be a banach spaces, $A \subset X$ open, $f, g : A \to Y$, $\lambda \in \mathbb{R}$ and $x_0 \in A$. If f and g are Frechet differentiable at x_0, then $f + \lambda g$ is Frechet differentiable at x_0 and $D(f+\lambda g)(x_0) = Df(x_0) + \lambda Dg(x_0)$.

Proof. Suppose that f and g are Frechet differentiable at x_0. Then

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + o(\|h\|) \quad \text{as } h \to 0$$

and

$$g(x_0 + h) = g(x_0) + Dg(x_0)(h) + o(\|h\|) \quad \text{as } h \to 0$$

This implies that

$$\begin{aligned}(f+\lambda g)(x_0+h) &= f(x_0+h) + \lambda g(x_0+h) \\ &= f(x_0) + Df(x_0)(h) + o(\|h\|) + \lambda g(x_0) + \lambda Dg(x_0)(h) + o(\|h\|) \\ &= (f+\lambda g)(x_0) + [Df(x_0) + \lambda Dg(x_0)](h) + o(\|h\|) \quad \text{as } h \to 0\end{aligned}$$

Since $Df(x_0)+\lambda Dg(x_0) \in L(X,Y)$, $f+\lambda g$ is Frechet differentiable at x_0 and $D(f+\lambda g)(x_0) = Df(x_0) + \lambda Dg(x_0)$. □

Exercise 6.2.5. Let X, Y be a banach spaces, $A \subset X$ open, $f : A \to Y$ and $x_0 \in A$. If f is Frechet differentiable at x_0, then f is Gateaux differentiable at x_0 and $df(x_0) = Df(x_0)$.

Proof. Suppose that f is Frechet differentiable at x_0. Then $f(x_0+h) = f(x_0)+Df(x_0)(h)+o(\|h\|)$ as $h \to 0$. Let $x \in X$. Then $f(x_0 + tx) - f(x_0) = tDf(x_0)(x) + o(t)$ as $t \to 0$. This implies that f is differentiable at x_0 in the direction x and

$$\begin{aligned}df(x_0)(x) &= \lim_{t\to 0} \frac{f(x_0+tx) - f(x_0)}{t} \\ &= Df(x_0)(x)\end{aligned}$$

Since $x \in X$ is arbitrary, f is Gateaux differentiable at x_0 and $df(x_0) = Df(x_0)$. □

Exercise 6.2.6. Let X be a Banach space, $A \subset X$ open, $f : A \to \mathbb{R}$ and $x_0 \in A$. If f is Frechet differentiable at x_0 and f has a local extremum at x_0, then $df(x_0) = 0$.

Proof. Suppose that f is Frechet differentiable at x_0 and f has a local extremum at x_0, then $df(x_0) = 0$. Two previous exercises imply that f is Gateaux differentiable at x_0 and

$$\begin{aligned} Df(x_0) &= df(x_0) \\ &= 0 \end{aligned}$$

□

Exercise 6.2.7. Chain Rule:
Let X, Y, Z be a Banach spaces, $A \subset X$ open, $B \subset Y$ open, $f : A \to Y$, $g : B \to Z$ and $x_0 \in A$. Suppose that $f(x_0) \in B$. If f is Frechet differentiable at x_0 and g is Frechet differentiable at $f(x_0)$, then $g \circ f$ is Frechet differentiable at x_0 and

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

Proof. Suppose that f is Frechet differentiable at x_0 and g is Frechet differentiable at $f(x_0)$. Then

(1) $f(x_0 + h) = f(x_0) + Df(x_0)(h) + o(\|h\|)$ as $h \to 0$
(2) $g(f(x_0) + k) = g(f(x_0)) + Dg(f(x_0))(k) + o(\|k\|)$ as $k \to 0$

□

Exercise 6.2.8. Let Y be a Banach space, $A \subset \mathbb{R}$ open and $f : A \to Y$. Then f is differentiable iff f is Frechet differentiable.

Proof. Suppose that f is Gateaux differentiable. Let $x_0 \in A$. A previous exercise implies that $df(x_0) \in L(\mathbb{R}, Y)$. By defintion,

$$\lim_{h \to 0} \left\| \frac{f(x_0 + h) - f(x_0)}{h} - df(x_0)(1) \right\| = 0$$

This is equivalent to saying that

$$f(x_0 + h) = f(x_0) + df(x_0)(h) + o(|h|) \qquad \text{as } h \to 0$$

So f is Frechet differentiable at x_0 and $Df(x_0) = df(x_0)$. □

6.3. **The Calc I Derivative.**

Definition 6.3.1. Calc I Derivative:
Let Y be a Banach space, $A \subset \mathbb{R}$ open, $f : A \to Y$ and $x_0 \in A$.

(1)
- If f is Frechet differentiable at x_0, we define the **calc I derivative of f at x_0**, denoted
$$f'(x_0) \text{ or } \frac{\mathrm{d}f}{\mathrm{d}t}(x_0)$$
by
$$\begin{aligned} f'(x_0) &= \lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t} \\ &= df(x_0)(1) \\ &= Df(x_0)(1) \end{aligned}$$
- If f is Frechet differentiable, we define $f' : A \to Y$ by $x \mapsto f'(x)$.

(2) Continuing inductively, we set $f^{(0)} = f$ and for $n \geq 2$,
- if $f^{(n-1)}$ is Frechet differentiable at x_0, we define the **(n)-th order calc I derivative of f at x_0**, denoted $f^{(n)}(x_0)$, by
$$f^{(n)} = [f^{(n-1)}]'$$
- if $f^{(n-1)}$ is Frechet differentiable, we define $f^{(n)} : A \to Y$ by
$$f^{(n)} = [f^{(n)}]'$$

Exercise 6.3.2. Let Y be a Banach space, $A \subset \mathbb{R}$ open and $f : A \to Y$. If f is n-th order Frechet differentiable, then for each $x_0 \in A$ and $k \in \{1, \cdots, n\}$,
$$f^{(k)}(x_0) = D^k f(x_0)(1^{\oplus k})$$

Proof. Let $x_0 \in A$. We proceed by induction. The base case is true by definition. Let $k \in \{1, \cdots, n\}$. Suppose the claim is true for $k-1$. Then
$$f^{(k-1)}(x_0) = D^{k-1} f(x_0)(1^{\oplus(k-1)})$$
Since f is n-th order Frechet differentiable,
$$D^{k-1} f(x_0 + h) = D^{k-1} f(x_0) + D^k f(x_0)(h) + o(\|h\|) \qquad \text{as } h \to 0$$
This implies that
$$\begin{aligned} f^{(k-1)}(x_0 + h) &= D^{k-1} f(x_0 + h)(1^{\oplus(k-1)}) \\ &= D^{k-1} f(x_0)(1^{\oplus(k-1)}) + D^k f(x_0)(h)(1^{\oplus(k-1)}) + o(\|h\|) \qquad \text{as } h \to 0 \end{aligned}$$
Therefore for each $h \in \mathbb{R}$,
$$Df^{(k-1)}(x_0)(h) = D^k f(x_0)(h)(1^{\oplus(k-1)})$$
and by definition,
$$\begin{aligned} f^{(k)}(x_0) &= [f^{(k-1)}]'(x_0) \\ &= Df^{(k-1)}(x_0)(1) \\ &= D^k f(x_0)(1^{\oplus k}) \end{aligned}$$
$\square$

Exercise 6.3.3. Let X, Y be Banach spaces, $A \subset X$ open, $f \in C^n(A, Y), x_0 \in A$, and $h \in X$. Suppose that $\{x_0 + tx : T \in [0,1]\} \subset A$. Define and $g : (0,1) \to Y$ by

$$g(t) = f(x_0 + th)$$

Then for each $k \in \{1 \dots, n\}$ and $t \in (0,1)$,

$$g^{(k)}(t) = D^k f(x_0 + th)(h^{\oplus k})$$

Proof. We proceed by induction. It is straightforward to show that the claim is true for $k = 1$.
Let $k \in \{1 \dots, n\}$. Suppose that $g^{(k-1)}(t) = D^{k-1} f(x_0 + th)(h^{\oplus(k-1)})$. Since $f \in C^k(A, Y)$,

$$D^{k-1} f(x_0 + s_0 h + th) = D^{k-1} f(x_0 + s_0 h) + D^k f(x_0 + s_0 h)(th) + o(\|t\|) \quad \text{as } t \to 0$$

The previous exercise implies that

$$\begin{aligned} g^{(k-1)}(s_0 + t) &= D^{k-1} g(s_0 + t)(1^{\oplus(k-1)}) \\ &= D^{k-1} f(x_0 + s_0 h + th)(h^{\oplus(k-1)}) \\ &= D^{k-1} f(x_0 + s_0 h)(h^{\oplus(k-1)}) + D^k f(x_0 + s_0 h)(th)(h^{\oplus(k-1)}) + o(\|t\|) \quad \text{as } t \to 0 \end{aligned}$$

Hence

$$Dg^{(k-1)}(s_0)(t) = D^k f(x_0 + s_0 h)(th)(h^{\oplus(k-1)})$$

and

$$\begin{aligned} g^{(k)}(t) &= Dg^{(k-1)}(t)(1) \\ &= D^k f(x_0 + th)(h^{\oplus k}) \end{aligned}$$

$\square$

6.4. **Taylor's Theorem.**

Note 6.4.1. This section makes use of the Bochner integral. For reference, see .

Exercise 6.4.1. Mean Value Theorem:
Let X, Y be a Banach spaces, $A \subset X$ open and convex and $f : A \to Y$. If f is Frechet differentiable, then for each $x, y \in A$, there exists $t \in (0,1)$ such that

$$\|f(x) - f(y)\| \leq \|Df(tx + (1-t)y)\|\|x - y\|$$

Hint: For $x, y \in A$ with $f(x) \neq f(y)$, using a Hahn-Banach argument, find $\lambda \in Y^*$ such that $\|\lambda\| = 1$ and $\lambda(f(x) - f(y)) = \|f(x) - f(y)\|$.

Proof. Suppose that f is Frechet differentiable. Let $x, y \in A$. The claim is clearly true when $f(x) = f(y)$. Suppose that $f(x) \neq f(y)$. An exercise in the section on linear functionals implies that there exists $\lambda \in Y^*$ such that $\lambda(f(x) - f(y)) = \|f(x) - f(y)\|$ and $\|\lambda\| = 1$ Define $g : [0,1] \to \mathbb{R}$ by

$$g(t) = \lambda(f(tx + (1-t)y))$$

Then g is continuous and (Frechet) differentiable on $(0,1)$ with

$$Dg(t)(h) = \lambda \circ Df(tx + (1-t)y)((x - y)h)$$

which implies that

$$\begin{aligned} g'(t) &= Dg(t)(1) \\ &= \lambda \circ Df(tx + (1-t)y)((x - y)) \end{aligned}$$

The mean value theorem implies that there exists $t \in (0,1)$ such that

$$\begin{aligned} \|f(x) - f(y)\| &= \lambda(f(x) - f(y)) \\ &= g(1) - g(0) \\ &= g'(t) \\ &= \lambda \circ Df(tx + (1-t)y)((x - y)) \end{aligned}$$

Taking absolute values, we see that

$$\begin{aligned} \|f(x) - f(y)\| &= |\lambda \circ Df(tx + (1-t)y)((x - y))| \\ &\leq \|\lambda\|\|Df(tx + (1-t)y)\|\|x - y\| \\ &\leq \|Df(tx + (1-t)y)\|\|x - y\| \end{aligned}$$

$\square$

Exercise 6.4.2. Let X, Y be a Banach spaces, $A \subset X$ open and convex and $f : A \to Y$. Suppose that f is Frechet differentiable. If for each $x \in A$, $Df(x) = 0$, then f is constant.

Proof. Suppose that for each $x \in A$, $Df(x) = 0$. Let $x, y \in A$. Then the mean value theorem implies that there exists $t \in (0,1)$ such that

$$\begin{aligned} \|f(x) - f(y)\| &\leq \|Df(tx + (1-t)y)\|\|x - y\| \\ &= 0 \end{aligned}$$

So $f(x) = f(y)$. $\square$

Exercise 6.4.3. Let X, Y be a Banach spaces, $A \subset X$ open and convex and $f, g : A \to Y$. Suppose that f and g are Frechet differentiable. If $Df = Dg$, then there exists $c \in Y$ such that $f = g + c$.

Proof. Suppose that $Df = Dg$. Then $D(f - g) = 0$ and the previous exercise implies that $f - g$ is constant. $\square$

Exercise 6.4.4. Let X, Y be a Banach spaces, $A \subset \mathbb{R}$ open and $f : A \to Y$. Suppose that f is Frechet differentiable. Then $f' \in C(A, Y)$ iff $f \in C^1(A, Y)$.

Proof. Suppose that $f' \in C(A, Y)$. Let $x, y \in A$ and $h \in \mathbb{R}$. Then

$$\begin{aligned}
\|(Df(x) - Df(y))(h)\| &= \|Df(x)(h) - Df(y)(h)\| \\
&= \|hf'(x) - hf'(y)\| \\
&= \|h(f'(x) - f'(y))\| \\
&= \|f'(x) - f'(y)\||h|
\end{aligned}$$

So $\|Df(x) - Df(y)\| \leq \|f'(x) - f'(y)\|$. Hence continuity of f' implies continuity of Df and $f \in C^1(A, Y)$. Conversely, suppose that $f \in C^1(A, Y)$. Let $x, y \in A$. Then

$$\begin{aligned}
\|f'(x) - f'(y)\| &= \|Df(x)(1) - Df(y)(1)\| \\
&= \|(Df(x) - Df(y))(1)\| \\
&\leq \|Df(x) - Df(y)\|
\end{aligned}$$

Hence continuity of Df implies continuity of f' and $f' \in C(A, Y)$. $\square$

Exercise 6.4.5. Let Y be a separable Banach space, $f : [a, b] \to Y$ continuous so that f is Bochner-integrable. Define $F : (a, b) \to Y$ by

$$F(x) = \int_{(a,x]} f dm$$

Then $F \in C^1((a, b), Y)$ and for each $x_0 \in (a, b)$ and $F'(x_0) = f(x_0)$.

Proof. Let $x_0 \in (a, b)$ and $h \in (0, b - x_0)$. Then continuity implies that

$$\begin{aligned}
\frac{1}{\|h\|}\left| \int_{(x_0,x_0+h]} f - f(x_0) dm \right| &\leq \frac{1}{\|h\|} \max_{x \in (x_0,x_0+h]} |f(x) - f(x_0)|\|h\| \\
&= \max_{x \in [x_0,x_0+h]} |f(x) - f(x_0)| \\
&\to 0 \text{ as } h \to 0
\end{aligned}$$

So

$$\int_{(x_0,x_0+h]} f - f(x_0) dm = o(\|h\|) \qquad \text{as } h \to 0$$

Therefore

$$\begin{aligned}F(x_0+h) &= \int_{(a,x_0+h]} f\,dm \\ &= \int_{(a,x_0]} f\,dm + \int_{(x_0,x_0+h]} f\,dm \\ &= \int_{(a,x_0]} f\,dm + hf(x_0) + \int_{(x_0,x_0+h]} f - f(x_0)dm \\ &= F(x_0) + hf(x_0) + o(\|h\|) \qquad \text{as } h \to 0\end{aligned}$$

The case is similar for $h \in (x_0 - b, 0)$. Since the map $h \mapsto f(x_0)h$ is bounded, F is Frechet differentiable at x_0 and $DF(x_0)(h) = f(x_0)h$. This implies that $F'(x_0) = f(x_0)$ and the previous exercise implies tells us that continuity of f implies continuity of DF. So $F \in C^1(A, Y)$. □

Exercise 6.4.6. Fundamental Theorem of Calculus: Let Y be a separable Banach space and $f \in C^1((a,b), Y)$. Then for each $x, x_0 \in (a,b)$, $x_0 < x$ implies that

(1) f' is Bochner integrable on $(x_0, x]$
(2)
$$f(x) - f(x_0) = \int_{(x_0,x]} f'dm$$

Proof. (1) Since $f \in C^1((a,b), Y)$, a previous exercise tells us that $f' \in C_Y(a,b)$. Let $x, x_0 \in (a,b)$. Suppose that $x_0 < x$. Choose $c, d \in (a,b)$ such that $a < c < x_0 < x < d < b$. Then f' is continuous on $[c,d]$ and hence Bochner-integrable on $(c,d]$ and $(x_0, x]$.

(2) Define $g : (c,d) \to Y$ by
$$g(\xi) = \int_{(c,\xi]} f'dm$$
Then the previous exercise implies that $g \in C^1_Y(c,d)$ and for each $t \in (c,d)$, $g'(t) = f'(t)$. Let $t \in (c,d)$ and $h \in \mathbb{R}$. Then
$$\begin{aligned}Dg(t)(h) &= hg'(t) \\ &= hf'(t) \\ &= Df(t)(h)\end{aligned}$$
So $Dg = Df$ on (c,d). A previous exercise implies that there exists $c \in Y$ such that $f = g + c$ on (c,d). Then
$$\begin{aligned}f(x) - f(x_0) &= g(x) + c - (g(x_0) + c) \\ &= g(x) - g(x_0) \\ &= \int_{(c,x]} f'dm - \int_{(c,x_0]} f'dm \\ &= \int_{(x_0,x]} f'dm\end{aligned}$$
□

Exercise 6.4.7. Let Y be a Banach space, $A \subset \mathbb{R}$ open and $g : A \to Y$. If g is n-th order Frechet differentiable, then

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{k=0}^{n-1} \frac{(1-t)^k}{k!} g^{(k)}(t) = \frac{(1-t)^{n-1}}{(n-1)!} g^{(n)}(t)$$

Proof. Taking the derivative yields a telescoping series. □

Exercise 6.4.8. Taylor's Theorem:
Let X be a Banach space, Y a separable Banach space, $A \subset X$ open and convex, $f \in C^n(A, Y)$, $x_0 \in A$, and $h \in X$. Suppose $x_0 + h \in A$. Then

$$f(x_0 + h) = \sum_{k=0}^{n-1} \frac{1}{k!} D^k f(x_0)(h^{\oplus k}) + R(x_0; h)$$

where

$$R(x_0; h) = \frac{1}{(n-1)!} \int_{(0,1)} (1-t)^{n-1} D^n f(x_0 + th)(h^{\oplus n}) dm(t)$$

Hint: Define $g : (0, 1) \to Y$ by

$$g(t) = f(x_0 + th)$$

Then use the previous exercise and the fundamental theorem of calculus.

Proof. For each $k \in \{1, \dots, n\}$, a previous exercise implies that $g^{(k)}(t) = D^k f(x_0 + th)(h^{\oplus k})$, so $g^{(k)}(0) = D^k f(x_0)(h^{\oplus k})$. The previous exercise and the fundamental theorem of calculus tell us that

$$\begin{aligned} f(x_0 + h) - \sum_{k=0}^{n-1} \frac{1}{k!} D^k f(x_0)(h^{\oplus k}) &= g(1) - \sum_{k=0}^{n-1} \frac{1}{k!} g^{(k)}(0) \\ &= \int_{(0,1)} \left[\frac{\mathrm{d}}{\mathrm{d}t} \sum_{k=0}^{n-1} \frac{(1-t)^k}{k!} g^{(k)}(t) \right] dm(t) \\ &= \int_{(0,1)} \frac{(1-t)^{n-1}}{(n-1)!} g^{(n)}(t) dm(t) \end{aligned}$$

□

6.5. **The Gradient.**

Definition 6.5.1. Let H be a Hilbert space, $f : H \to \mathbb{R}$ and $x_0 \in H$. Suppose that f is Frechet differentiable at x_0. Then $Df(x_0) \in H^*$. We define the **gradient of** f **at** x_0, denoted $\nabla f(x_0) \in H$, via the Riesz representation theorem to be the unique element of H satisfying

$$\langle \nabla f(x_0), y \rangle = Df(x_0)(y) \quad \text{for each } y \in H$$

7. Convexity

7.1. Introduction.

Note 7.1.1. In this section, we assume all vector spaces are real.

Definition 7.1.1. Let X be a vector space and $A \subset X$. Then A is said to be **convex** if for each $x, y \in A$, and $t \in [0,1]$, $tx + (1-t)y \in A$.

Definition 7.1.2. Let X be a vector space and $f : A \to R$. Then f is said to be **convex** if for each $x, y \in A$, $t \in [0,1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Exercise 7.1.3. Let X be a Banach space, $A \subset X$ open and convex, and $f \in C^2(A)$. Then f is convex (resp. strictly convex) iff for each $x \in A$, $D^2 f(x)$ is positive semidefinite (resp. positive definite).
Hint: For $x, y \in X$, consider the function $g : (0,1) \to \mathbb{R}$ given by $g(t) = f(tx + (1-t)y)$

Proof. □

Exercise 7.1.4. Let X be a vector space, $f \in X^*$ and $g : X \to \mathbb{R}$ constant. Then f and g are convex.

Proof. Let $x, y \in X$ and $t \in [0,1]$. Put $c = g(0)$. Then

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y)$$

and

$$\begin{aligned} g(tx + (1-t)y) &= c \\ &= tc + (1-t)c \\ &= tg(x) + (1-t)g(y) \end{aligned}$$

So f and g are convex. □

Exercise 7.1.5. Let X be a vector space, $A \subset X$ convex, $f, g : A \to \mathbb{R}$ and $\lambda \geq 0$. If f, g are convex, then

(1) $f + g$ is convex
(2) λf is convex

Proof. Suppose that f and g are convex. Let $x, y \in A$ and $t \in [0,1]$. Then

$$\begin{aligned} (f + \lambda g)(tx + (1-t)y) &= f(tx + (1-t)y) + \lambda g(tx + (1-t)y) \\ &\leq tf(x) + (1-t)f(y) + t\lambda g(x) + (1-t)\lambda g(y) \\ &= t(f(x) + \lambda g(x)) + (1-t)(f(y) + \lambda g(y)) \\ &= t(f + \lambda g)(x) + (1-t)(f + \lambda g)(y) \end{aligned}$$

□

Definition 7.1.6. Let X be a vector space and $f : X \to \mathbb{R}$. Then f is said to be **affine** if there exists $\phi \in X^*$, $a \in R$ constant such that $f = \phi + a$.

Exercise 7.1.7. Let X be a vector space and $f : X \to \mathbb{R}$. If f is affine, then f is convex.

Proof. Suppose that f is affine. Then there exists $\phi \in X^*$, $a \in R$ constant such that $f = \phi + a$. Then ϕ is convex and $g : X \to \mathbb{R}$ defined by $g(x) = a$ is convex. So $f = \phi + g$ is convex. □

Exercise 7.1.8. Let X be a vector space, $A \subset X$ convex, $f : \mathbb{R} \to \mathbb{R}$ and $g : A \to \mathbb{R}$. If f is convex and increasing and g is convex, then $f \circ g$ is convex.

Proof. Let $t \in [0, 1]$ and $x, y \in A$. Then convexity of g implies that

$$g(tx + (1-t)y) \leq tg(x) + (1-t)g(y)$$

and we have

$$\begin{aligned} f \circ g(tx + (1-t)y) &= f(g(tx + (1-t)y)) \\ &\leq f(tg(x) + (1-t)g(y)) && (f \text{ increasing}) \\ &\leq tf(g(x)) + (1-t)f(g(y)) && (f \text{ convex}) \\ &= tf \circ g(x) + (1-t)f \circ g(y) \end{aligned}$$

So $f \circ g$ is convex. □

Exercise 7.1.9. Let X be a vector space, $A \subset X$ convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. Then f has a local minimum point at x_0 iff f has a global minimum point at x_0.

Proof. If f has a global minimum point at x_0, then f has a local minimum point at x_0. Conversely, suppose that f has a local minimum point at x_0. Then there exists $\delta > 0$ such that for each $x \in B(x_0, \delta)$, $f(x_0) \leq f(x)$. For the sake of contradiction, suppose that f does not have a global minimum point at x_0. Then there exits $x' \in A$ such that $f(x') < f(x_0)$. Put $t_0 = \min(\frac{\delta}{\|x' - x_0\| + 1}, 1) > 0$. Let $t \in (0, t_0)$, then

$$\begin{aligned} \|(tx' + (1-t)x_0) - x_0\| &= t\|x' - x_0\| \\ &< \frac{\|x' - x_0\|\delta}{\|x' - x_0\| + 1} \\ &< \delta \end{aligned}$$

so that $tx' + (1-t)x_0 \in B(x_0, \delta)$ and hence $f(x_0) \leq f(tx' + (1-t)x_0)$. Therefore

$$\begin{aligned} f(x_0) &\leq f(tx' + (1-t)x_0) \\ &\leq tf(x') + (1-t)f(x_0) \quad \text{(convexity of } f) \\ &< tf(x_0) + (1-t)f(x_0) \\ &= f(x_0) \end{aligned}$$

which is a contradiction. Hence f has a global minimum point at x_0. □

Definition 7.1.10. Let X, Y be vector spaces, $A \subset X \oplus Y$. For $y \in Y$, define

$$A^y = \{x \in X : (x, y) \in A\}$$

and $f^y : A^y \to \mathbb{R}$ by

$$f^y(x) = f(x, y)$$

Exercise 7.1.11. Let X, Y be vector spaces, $A \subset X \oplus Y$ convex and $f : A \to \mathbb{R}$ convex. Then for each $y \in \pi_2(A)$,

(1) A^y is convex

(2) f^y is convex

where $\pi_2 : X \times Y \to Y$, the canonical projection of $X \times Y$ onto Y given by $\pi_2(x, y) = y$.

Proof. Let $y \in \pi_2(A)$, $x_1, x_2 \in A^y$ and $t \in [0,1]$. Then by definition, (x_1, y), $(x_2, y) \in A$.

(1) Convexity of A implies that $(tx_1 + (1-t)x_2, y) \in A$. Hence $tx_1 + (1-t)x_2 \in A^y$ and A^y is convex.

(2) Convexity of f implies that

$$\begin{aligned} f^y(tx_1 + (1-t)x_2) &= f(tx_1 + (1-t)x_2, y) \\ &= f(t(x_1, y) + (1-t)(x_2, y)) \\ &\leq tf(x_1, y) + (1-t)f(x_2, y) \\ &= tf^y(x_1) + (t-t)f^y(x_2) \end{aligned}$$

and so f^y is convex.

□

Exercise 7.1.12. Let X, Y be vector spaces and $A \subset X, B \subset Y$. If A and B are convex, then $A \times B \subset X \oplus Y$ is convex.

Proof. Suppose that A and B are convex. Let $(x_1, y_1), (x_2, y_2) \in A \times B$ and $t \in [0,1]$. Convexity of A and B implies that $tx_1 + (1-t)x_2 \in A$ and $ty_1 + (1-t)y_2 \in B$. Therefore

$$\begin{aligned} t(x_1, y_1) + (1-t)(x_2, y_2) &= (tx_1 + (1-t)x_2, ty_1 + (1-t)y_2) \\ &\in A \times B \end{aligned}$$

□

Exercise 7.1.13. Let X, Y be vector spaces and $A \subset X$, $B \subset Y$ convex (implying that $A \times B$ is convex) and $f : A \times B \to \mathbb{R}$ convex. Suppose that for each $y \in B$, $\{f(x, y) : x \in A\}$ is bounded below. Then $\inf\limits_{y \in B} f^y$ is convex

Proof. Put $g = \inf\limits_{y \in B} f^y$. Let $x_1, x_2 \in A$, $y_1, y_2 \in B$ and $t \in [0,1]$. Put $y' = ty_1 + (1-t)y_2$. Then convexity of f implies that

$$\begin{aligned} g(tx_1 + (1-t)x_2) &\leq f^{y'}(tx_1 + (1-t)x_2) \\ &= f(tx_1 + (1-t)x_2, ty_1 + (1-t)y_2) \\ &= f(t(x_1, y_1) + (1-t)(x_2, y_2)) \\ &\leq tf(x_1, y_1) + (1-t)f(x_2, y_2) \\ &= tf^{y_1}(x_1) + (1-t)f^{y_2}(x_2) \end{aligned}$$

Since $y_1 \in B$ is arbitrary, we have that

$$g(tx_1 + (1-t)x_2) \leq tg(x_1) + (1-t)f^{y_2}(x_2)$$

Similarly, since $y_2 \in B$ is arbitrary, we have that

$$g(tx_1 + (1-t)x_2) \leq tg(x_1) + (1-t)g(x_2)$$

and f is convex. □

Exercise 7.1.14. Let X be a vector space, $A \subset X$ convex and $(f_\lambda)_{\lambda \in \Lambda} \subset \mathbb{R}^A$. Suppose that for each $\lambda \in \Lambda$, f_λ is convex. Then $\sup\limits_{\lambda \in \Lambda} f_\lambda$ is convex.

Proof. Define $f = \sup\limits_{\lambda \in \Lambda} f_\lambda$. Let $x, y \in A$, $t \in [0,1]$ and $\lambda \in \Lambda$. Then

$$\begin{aligned} f_\lambda(tx + (1-t)y) &\le t f_\lambda(x) + (1-t) f_\lambda(y) \\ &\le t f(x) + (1-t) f(y) \end{aligned}$$

Since $\lambda \in \Lambda$ is arbitrary, $f(tx + (1-t)y) \le t f(x) + (1-t) f(y)$. □

Exercise 7.1.15. Let X be a normed vector space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. If f is continuous at x_0, then f is locally Lipschitz at x_0.
Hint: Given x_1, x_2 near x_0 Choose a z near x_0 s.t. x_1 is a convex combination of x_2 and z. Then repeat but with x_2 as a convex combination of x_1 and z

Proof. By continuity, f is locally bounded at x_0. So there exist $M, \delta > 0$ such that $B(x_0, \delta) \subset A$ and for each $x \in B(x_0, \delta)$, $|f(x)| \le M$. Put $\delta' = \frac{\delta}{2}$ and choose $U = B(x_0, \delta')$. Then $U \subset A$, U is open and $U \in N_{x_0}$.
Let $x_1, x_2 \in U$. Suppose that $x_1 \ne x_2$. Define $\alpha = \|x_1 - x_2\| > 0$, $p = \frac{\alpha}{\alpha + \delta'}$, $q = 1 - p$ and $z = p^{-1}(x_1 - q x_2)$. Then $x_1 = pz + q x_2$ and

$$\begin{aligned} \|z - x_1\| &= \|(p^{-1} - 1)x_1 - p^{-1} q x_2\| \\ &= \frac{1-p}{p}\alpha \\ &= \frac{\delta'}{\alpha}\alpha \\ &= \delta' \end{aligned}$$

Therefore

$$\begin{aligned} \|z - x_0\| &\le \|z - x_1\| + \|x_1 - x_0\| \\ &< \delta' + \delta' \\ &= \delta \end{aligned}$$

So $z \in B(x_0, \delta)$, which implies that

$$\begin{aligned} f(z) - f(x_2) &\le |f(z) - f(x_2)| \\ &\le |f(z)| + |f(x_2)| \\ &\le 2M \end{aligned}$$

Since $x_1 = pz + q x_2$, convexity of f implies that $f(x_1) \le p f(z) + q f(x_2)$. Hence

$$\begin{aligned} f(x_1) - f(x_2) &\le p f(z) - p f(x_2) \\ &= p(f(z) - f(x_2)) \\ &\le p 2M \\ &= \frac{\alpha}{\alpha + \delta'} 2M \\ &\le \alpha 2M \\ &= 2M \|x_1 - x_2\| \end{aligned}$$

Similarly, choosing $z = p^{-1}(x_2 - qx_1)$, yields $f(x_2) - f(x_1) \leq 2M\|x_1 - x_2\|$ which implies that

$$|f(x_1) - f(x_2)| \leq 2M\|x_1 - x_2\|$$

and f is Lipschitz on U. $\square$

7.2. **The Subdifferential.**

Exercise 7.2.1. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex, $x_0 \in A$ and $x \in X$. Define $T = \{t \in \mathbb{R} : x_0 + tx \in A\}$. Then there exist $a, b \in (0, \infty]$ such that $T = (-a, b)$.

Proof. Continuity of scalar multiplication and addition implies that T is an open neighborhood of 0. Let $t > 0$ and $s \in [0, t]$. Then $\frac{s}{t} \in [0, 1]$ and by convexity of A, $x_0 + tx \in A$ implies that

$$\begin{aligned} x_0 + sx &= \frac{s}{t}(x_0 + tx) + \left(1 - \frac{s}{t}\right)x_0 \\ &\in A \end{aligned}$$

Thus $[0, t] \subset T$. Similarly, $x_0 - tx \in A$ implies that $[-t, 0] \subset T$.
Define $a, b \in (0, \infty]$ by $a = \sup\{t > 0 : x_0 - tx \in A\}$ and $b = \sup\{t > 0 : x_0 + tx \in A\}$. Then $(-a, b) = T$. □

Definition 7.2.2. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex, $x_0 \in A$ and $x \in X$. Define T as in the previous exercise and choose $t_0 > 0$ such that $(-t_0, t_0) \subset T$. For $t \in (0, t_0)$, define the difference quotient $q : (-t_0, t_0) \setminus \{0\} \to \mathbb{R}$ by

$$q(t) = \frac{f(x_0 + tx) - f(x_0)}{t}$$

Exercise 7.2.3. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex, $x_0 \in A$ and $x \in X$. Define t_0 as above. Then

(1) $q(t)$ is increasing on $(0, t_0)$
(2) $q(-t)$ decreasing on $(0, t_0)$

Hint: As an example, look at the graph of $f(x) = x^2$. For the algebra, start at the desired end inequality and work backwards

Proof. Let $s, t \in (0, t_0)$ and suppose that $s \leq t$. Then $x_0 + sx$, $x_0 + tx \in A$. Note that since $0 < s \leq t$, $\frac{s}{t} \in (0, 1]$ and $1 - \frac{s}{t} = \frac{t-s}{t} \in (0, 1]$. Also, since A is convex, we have that

$$\left(\frac{t-s}{t}\right)x_0 + \left(\frac{s}{t}\right)(x_0 + tx) \in A$$

Convexity of f implies that

$$\begin{aligned} f(x_0 + sx) &= f\left(\left(\frac{t-s}{t}\right)x_0 + \left(\frac{s}{t}\right)(x_0 + tx)\right) \\ &\leq \left(\frac{t-s}{t}\right)f(x_0) + \left(\frac{s}{t}\right)f(x_0 + tx) \end{aligned}$$

This implies that

$$tf(x_0 + sx) \leq (t - s)f(x_0) + sf(x_0 + tx)$$

and after rearranging, we get

$$tf(x_0 + sx) - tf(x_0) \leq sf(x_0 + tx) - sf(x_0)$$

and so finally, dividing both sides by st, we obtain

$$\begin{aligned} q(s) &= \frac{f(x_0 + sx) - f(x_0)}{s} \\ &\leq \frac{f(x_0 + tx) - f(x_0)}{t} \\ &= q(t) \end{aligned}$$

as desired.

Similar to (1).

□

Exercise 7.2.4. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex, $x_0 \in A$ and $x \in X$. Define t_0 as before. Then for each $t \in (0, t_0)$,

$$q(-t) \leq q(t)$$

Hint: for sufficiently small t, convexity of f implies that $f(x_0) \leq \frac{1}{2}f(x_0 - 2tx) + \frac{1}{2}f(x_0 + 2tx)$

(1) *Proof.* Choose t_0 as in the previous exercise. Since convexity of f implies that for each $t \in (0, t_0/2)$,

$$f(x_0) \leq \frac{1}{2}f(x_0 - 2tx) + \frac{1}{2}f(x_0 + 2tx)$$

we have that for each $t \in (0, t_0/2)$,

$$\begin{aligned} q(-2t) &= \frac{f(x_0 - 2tx) - f(x_0)}{-2t} \\ &\leq \frac{f(x_0 + 2tx) - f(x_0)}{2t} \\ &= q(2t) \end{aligned}$$

So for each $t \in (0, t_0)$, $q(-t) \leq q(t)$. □

Exercise 7.2.5. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. Then

(1) f is left-hand and right-hand Gateaux differentiable at x_0 with $d^- f(x_0) \leq d^+ f(x_0)$
(2) for each $x \in X$, $d^- f(x_0)(x) = -d^+ f(x_0)(-x)$

Proof.

(1) Let $x \in X$. Choose $t_0 > 0$ as in the previous two exercises. Let $t, u \in (0, t_0)$. Choose $s \in (0, \min(u, t))$. The previous two exercises imply that

$$\begin{aligned} q(-u) &\leq q(-s) \\ &\leq q(s) \\ &\leq q(t) \end{aligned}$$

and therefore $q(t)$ is an upper bound for $\{q(-u) : u \in (0, t_0)\}$ and $d^- f(x_0)(x) = \sup\limits_{u \in (0, t_0)} q(-u)$ exists with $d^- f(x_0)(x) \leq q(t)$.

Since $t \in (0, t_0)$ is arbitrary, $d^- f(x_0)(x)$ is a lower bound for $\{q(t) : t \in (0, t_0)\}$. Therefore

$$d^+ f(x_0)(x) = \inf_{t \in (0, t_0)} q(t)$$

exists with $d^+ f(x_0)(x) \geq d^- f(x_0)(x)$.

(2) By definition, we have

$$\begin{aligned} d^- f(x_0)(x) &= \lim_{t \to 0^+} \frac{f(x_0 + -tx) - f(x_0)}{-t} \\ &= -\lim_{t \to 0^+} \frac{f(x_0 + -tx) - f(x_0)}{t} \\ &= -d^+ f(x_0)(-x) \end{aligned}$$

□

Exercise 7.2.6. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. Then $d^+ f(x_0) : X \to \mathbb{R}$ is a sublinear functional.

Proof. Let $x, y \in X$ and $k \geq 0$. If $k = 0$, then clearly

$$d^+ f(x_0)(kx) = kd^+(x_0)(x)$$

If $k > 0$. Then

$$\begin{aligned} d^+ f(x_0)(kx) &= \lim_{t \to 0^+} \frac{f(x_0 + tkx) - f(x_0)}{t} \\ &= k \lim_{t \to 0^+} \frac{f(x_0 + tkx) - f(x_0)}{tk} \\ &= kd^+ f(x_0)(x) \end{aligned}$$

Define $t_0 > 0$ as before and let $t \in (0, \frac{t_0}{2})$. Note that

$$x_0 + tx + ty = \frac{1}{2}(x_0 + 2tx) + \frac{1}{2}(x_0 + 2ty)$$

Convexity of f implies that

$$f(x_0 + tx + ty) \leq \frac{1}{2} f(x_0 + 2tx) + \frac{1}{2} f(x_0 + 2ty)$$

which implies that

$$\frac{f(x_0 + tx + ty) - f(x_0)}{t} \leq \frac{f(x_0 + 2tx) - f(x_0)}{2t} + \frac{f(x_0 + 2ty) - f(x_0)}{2t}$$

Therefore

$$\begin{aligned}
d^+f(x_0)(x+y) &= \lim_{t\to 0^+} \frac{f(x_0+t(x+y))-f(x_0)}{t} \\
&= \lim_{t\to 0^+} \frac{f(x_0+tx+ty)-f(x_0)}{t} \\
&\leq \lim_{t\to 0^+} \left[\frac{f(x_0+2tx)-f(x_0)}{2t} + \frac{f(x_0+2ty)-f(x_0)}{2t}\right] \\
&= \lim_{t\to 0^+} \frac{f(x_0+2tx)-f(x_0)}{2t} + \lim_{t\to 0^+} \frac{f(x_0+2ty)-f(x_0)}{2t} \\
&= d^+f(x_0)(x) + d^+f(x_0)(y)
\end{aligned}$$

□

Exercise 7.2.7. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. Then for each $x \in A$,

$$d^+f(x_0)(x-x_0) \leq f(x) - f(x_0)$$

Proof. Let $x \in A$. Define $T = \{t \in \mathbb{R} : x_0 + t(x - x_0) \in A\}$ similarly to earlier. Clearly $1 \in T$ and

$$\begin{aligned}
d^+f(x_0)(x-x_0) &= \inf_{t\in(0,1]} \frac{f(x_0+t(x-x_0))-f(x_0)}{t} \\
&\leq f(x) - f(x_0)
\end{aligned}$$

□

Exercise 7.2.8. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. If f is continuous at x_0, then $d^+f(x_0)$ is Lipschitz (equivalently bounded).

Proof. Suppose that f is continuous at x_0. A previous exercise about convex functions tells us that f is locally Lipschitz at x_0, so there exists $\delta, M > 0$ such that for each $x_1, x_2 \in B(x_0, \delta)$, $|f(x_1) - f(x_2)| \leq M\|x_1 - x_2\|$. Let $x \in X$ and define $t_0 = \frac{\delta}{\|x\|+1}$ so that for each $t \in (0, t_0)$,

$$\begin{aligned}
\|(x_0+tx)-x_0\| &= t\|x\| \\
&\leq t_0\|x\| \\
&= \frac{\delta\|x\|}{\|x\|+1} \\
&< \delta
\end{aligned}$$

and $x_0 + tx \in B(x_0, \delta)$. Then for each $t \in (0, t_0)$,

$$\begin{aligned}
d^+f(x_0)(x) &\leq \frac{f(x_0+tx)-f(x_0)}{t} \\
&\leq \frac{|f(x_0+tx)-f(x_0)|}{t} \\
&\leq t^{-1}M\|(x_0+tx)-x_0\| \\
&= M\|x\|
\end{aligned}$$

Thus $d^+f(x_0)$ is a bounded sublinear functional and a previous exercise in the section on sublinear functionals implies this is eqivalent to $d^+f(x_0)$ being Lipschitz. □

Exercise 7.2.9. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. If f is continuous at x_0, then there exists $\phi \in X^*$ such that $\phi \leq d^+ f(x_0)$.

Proof. Suppose that f is continuous at x_0. The previous exercise implies that $d^+ f(x_0)$ is Lipschitz (equivalently bounded). A previous exercise in the section discussing sublinear functionals tells us that boundedness of $d^+ f(x_0)$ implies that there exists $\phi \in X^*$ such that $\phi \leq d^+ f(x_0)$. □

Definition 7.2.10. Subdifferential:
Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. We define the **subdifferential of f at** x_0, denoted $\partial f(x_0)$, to be

$$\partial f(x_0) = \{\phi \in X^* : \text{for each } x \in A, f(x_0) + \phi(x - x_0) \leq f(x)\}$$

Exercise 7.2.11. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. If f is continuous at x_0, then $\partial f(x_0) \neq \varnothing$.

Proof. Suppose that f is continuous at x_0. The previous exercise tells us that there exists $\phi \in X^*$ such that $\phi \leq d^+ f(x_0)$. Let $x \in A$. A previous exercise implies that

$$\begin{aligned} \phi(x - x_0) &\leq d^+ f(x_0)(x - x_0) \\ &\leq f(x) - f(x_0) \end{aligned}$$

Then $f(x_0) + \phi(x - x_0) \leq f(x)$. □

Exercise 7.2.12. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex, $\phi \in X^*$ and $x_0 \in A$. Then

(1) for each $x \in A$,
$$\phi(x - x_0) \leq f(x) - f(x_0)$$
iff
$$\phi \leq d^+ f(x_0)$$
(2) $\partial f(x_0) = \{\phi \in X^* : \phi \leq d^+ f(x_0)\}$

Proof.

(1) Suppose that for each $x \in A$, $\phi(x - x_0) \leq f(x) - f(x_0)$. Let $x \in X$. Define t_0 as before. Then for each $t \in (0, t_0)$,
$$\begin{aligned} t\phi(x) &= \phi((x_0 + tx) - x_0) \\ &\leq f(x_0 + tx) - f(x_0) \end{aligned}$$
This implies that $\phi(x) \leq d^+ f(x_0)(x)$.
Conversely, suppose that $\phi \leq d^+ f(x_0)$. Let $x \in A$. A previous exercise implies that,
$$\begin{aligned} \phi(x - x_0) &\leq d^+ f(x_0)(x - x_0) \\ &\leq f(x) - f(x_0) \end{aligned}$$
(2) Clear.

□

Exercise 7.2.13. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. If f is continuous at x_0, then the following are equivalent:

(1) f is Gateaux differentiable at x_0

(2) $d^+f(x_0)$ is linear
(3) $\#\partial f(x_0) = 1$

Proof. Suppose that f is continuous at x_0. Then $d^+f(x_0)$ is Lipschitz and bounded.

- (1) $\Rightarrow$ (2):
  Suppose that f is Gateaux differentiable at x_0. Let $x \in X$. Then a previous exercise implies that
  $$\begin{aligned} -df^+(x_0)(-x) &= df^-f(x_0)(x) \\ &= df^+f(x_0)(x) \end{aligned}$$
  An exercise in the section on sublinear functionals implies that $df^+f(x_0)$ is linear.
- (2) $\Rightarrow$ (3):
  Suppose that $df^+f(x_0)$ is linear. Let $\phi \in \partial f(x_0)$. The previous exercise implies that $\phi \leq df^+f(x_0)$. Equivalence of linearity in the section on sublinear functionals implies that $d^+f(x_0) = \phi$.
- (3) $\Rightarrow$ (1):
  Suppose that $\#\partial f(x_0) = 1$. Since $\partial f(x_0) = \{\phi \in X^* : \phi \leq d^+f(x_0)\}$, equivalence of linearity in the section on sublinear functionals implies that $d^+f(x_0)$ is linear. This implies that $d^+f(x_0) = d^-f(x_0)$ and which implies that f is Gateaux differentiable at x_0.

$\square$

Exercise 7.2.14. Let X be a Banach space, $A \subset X$ open and convex, $f : A \to \mathbb{R}$ convex and $x_0 \in A$. If f is continuous at x_0, then f has a global minimum point at x_0 iff $0 \in \partial f(x_0)$.

Proof. Suppose that f has a global minimum point at x_0 iff $0 \in \partial f(x_0)$ Let $x \in X$. Then
$$\begin{aligned} d^+f(x_0)(x) &= \lim_{t \to 0^+} \frac{f(x_0 + tx) - f(x_0)}{t} \\ &\geq 0 \end{aligned}$$
So $0 \leq df^+(x_0)$ and $0 \in \partial f(x_0)$.
Conversely, suppose that $0 \in \partial f(x_0)$. Let $x \in A$. Then
$$\begin{aligned} 0 &= 0(x - x_0) \\ &\leq f(x) - f(x_0) \end{aligned}$$
So that $f(x_0) \leq f(x)$ which implies that f has a global minimum point at x_0. $\square$

7.3. **Conjugacy.**

Definition 7.3.1. Let X be a Banach space, $A \subset X$ and $f : A \to \mathbb{R}$. Define $A^* \subset X^*$ and $f^* : A^* \to \mathbb{R}$ by

$$A^* = \left\{\phi \in X^* : \sup_{x \in A} \left[\phi(x) - f(x)\right] < \infty\right\}$$

and

$$f^*(\phi) = \sup_{x \in A} \left[\phi(x) - f(x)\right]$$

If X is a Hilbert space, we may define $A^* \subset X$ and $f^* : A^* \to \mathbb{R}$ via the Riesz representation theorem by

$$A^* = \left\{y \in X : \sup_{x \in A} \left[\langle y, x\rangle - f(x)\right] < \infty\right\}$$

and $f^* : A^* \to \mathbb{R}$ and

$$f^*(y) = \sup_{x \in A} \left[\langle y, x\rangle - f(x)\right]$$

Exercise 7.3.2. Let X be a Banach space, $A \subset X$ and $f : A \to \mathbb{R}$. Then f^* is convex.

Proof. For $x \in A$, define $g_x : X^* \to [\infty, \infty)$ by $g_x(\phi) = \phi(x) - f(x)$. Then for each $x \in A$, g_x is convex since it is affine. Thus $f^* = \sup\limits_{x \in A} g_x$ is convex. □

Exercise 7.3.3. Let X be a Banach space, $A \subset X$ and $f : A \to \mathbb{R}$. Then for each $x \in X$ and $\phi \in X^*$, $f(x) \geq \phi(x) - f^*(\phi)$.

Proof. Clear □

Exercise 7.3.4.

Definition 7.3.5. Let

Definition 7.3.6. ∂f

Exercise 7.3.7.

7.4. **Functional Optimization.**

Exercise 7.4.1. Let X be a Banach space, $(S, \mathcal{S}, \mu)$ a measure space, $A \subset X$, $K \in L^0(A, \mathbb{R})$ and $\Lambda \subset L^0(S, A) \cap \{f : S \to A : K \circ f \in L^1(\mu)\}$. Suppose that A and Λ are convex. Define $\phi : \Lambda \to \mathbb{R}$ by

$$\phi f = \int K \circ f d\mu$$

Then K is convex implies that ϕ is convex.

Proof. Suppose that K is convex. Let $t \in [0, 1]$ and $f, g \in \Lambda$. Convexity of K implies that for each $s \in S$,

$$K[tf(s) + (1 - t)g(s)] \leq tK[f(s)] + (1 - t)K[g(s)]$$

So

$$K \circ [tf + (1 - t)g] \leq tK \circ f + (1 - t)K \circ g$$

Therefore

$$\begin{aligned} \phi[tf + (1 - t)g] &= \int K \circ [tf + (1 - t)g] d\mu \\ &\leq \int tK \circ f + (1 - t)K \circ g d\mu \\ &= t \int K \circ f d\mu + (1 - t) \int K \circ g d\mu \\ &= t\phi f + (1 - t)\phi g \end{aligned}$$

and ϕ is convex. $\square$

8. Appendix

8.1. Summation.

Definition 8.1.1. Let $f : X \to [0, \infty)$, Then we define

$$\sum_{x \in X} f(x) := \sup_{\substack{F \subset X \\ F \text{ finite}}} \sum_{x \in F} f(x)$$

This definition coincides with the usual notion of summation when X is countable. For $f : X \to \mathbb{C}$, we can write $f = g + ih$ where $g, h : X \to \mathbb{R}$. If

$$\sum_{x \in X} |f(x)| < \infty,$$

then the same is true for g^+, g^-, h^+, h^-. In this case, we may define

$$\sum_{x \in X} f(x)$$

in the obvious way.

The following note justifies the notation $\sum_{x \in X} f(x)$ where $f : X \to \mathbb{C}$.

Note 8.1.1. Let $f : X \to \mathbb{C}$ and $\alpha : X \to X$ a bijection. If $\sum_{x \in X} |f(x)| < \infty$, then $\sum_{x \in X} f(\alpha(x)) = \sum_{x \in X} f(x)$.

8.2. **Asymptotic Notation.**

Definition 8.2.1. Let X be a topological space, Y, Z be normed vector spaces, $f : X \to Y$, $g : X \to Z$ and $x_0 \in X \cup \{\infty\}$. Then we write

$$f = o(g) \qquad \text{as } x \to x_0$$

if for each $\epsilon > 0$, there exists $U \in \mathcal{N}_{x_0}$ such that U is open and for each $x \in U$,

$$\|f(x)\| \leq \epsilon \|g(x)\|$$

Exercise 8.2.2. Let X be a topological space, Y, Z be normed vector spaces, $f : X \to Y$, $g : X \to Z$ and $x_0 \in X \cup \{\infty\}$. If there exists $U \in \mathcal{N}_{x_0}$ such that U is open and for each $x \in U \setminus \{x_0\}$, $g(x) > 0$, then

$$f = o(g) \text{ as } x \to x_0 \quad \text{iff} \quad \lim_{x \to x_0} \frac{\|f(x)\|}{\|g(x)\|} = 0$$