BRAC UNIVERSITY

Inspiring Excellence

Functional Analysis (MAT433)

Lecture Notes

Preface

This series of lecture notes has been prepared for aiding students who took the BRAC University course **Functional Analysis (MAT433)** in Fall 2024 semester. These notes were typeset under the supervision of mathematician **Dr. Syed Hasibul Hassan Chowdhury**. The main goal of this typeset is to have an organized digital version of the notes, which is easier to share and handle. If you see any mistakes or typos, please send an email at atonuroychowdhury@gmail.com

The typsetting team comprises of Atonu Roy Chowdhury, Tanmoy Kumar Sarkar, Md Mubasshir Chowdhury, Hamidul Muslim Miad, A F Mashrafi.

Contents

1 Recap of Real Analysis, Topology and Linear Algebra

§1.1 Metric Spaces

Definition 1.1 (Metric Space). A metric space is a pair (X, d) where X is a set and d is a metric on X satisfying the following properties:

- (M1) $0 \leq d(x, y) < \infty$
- (M2) $d(x, y) = 0 \iff x = y$
- (M3) $d(x, y) = d(y, x)$
- (M4) $d(x, y) \leq d(x, z) + d(z, y)$

Example 1.1 (Euclean Space $\mathbb{R}^n$). The metric space $(\mathbb{R}^n, d_{\text{Euc}})$ consists of the set $\mathbb{R}^n$, the n dimensional Euclidean space, which contains all ordered n-tuples of real numbers, written as

$$x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n)$$

And the metric d_{Euc} is defined as:

$$d_{\text{Euc}}(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2)}$$

Example 1.2 (Unitary space $\mathbb{C}^n$). The metric space $(\mathbb{C}^n, d_{\text{uni}})$ consists of the set $\mathbb{C}^n$, which contains all ordered n-tuples of complex numbers, written as

$$z = (\alpha_1, \alpha_2, \ldots, \alpha_n), w = (\beta_1, \beta_2, \ldots, \beta_n)$$

And the metric d_{uni} is defined as:

$$d_{\text{uni}}(z, w) = \sqrt{\sum_{i=1}^{n}|\alpha_i - \beta_i|^2}$$

$\mathbb{C}^n$ is called the n-dimensional unitary space. It is sometimes called the complex Euclidean n-space.

Example 1.3 (Sequence Space). The metric space consists of the set of all bounded sequences of complex numbers, i.e every element of the set is a complex sequence,

$$x = (x_1, x_2, \ldots) \text{ briefly written as } x = \{x_i\}$$
$$y = (y_1, y_2, \ldots) \text{ briefly written as } y = \{y_i\}$$

Such that for all $i \in \mathbb{N}$, $|x_i| \leq M_x$ for some $M_x \in \mathbb{R}$depending on x but not on i . The metric d is defined as:

$$d(x, y) = \sup_{i \in \mathbb{N}} |x_i - y_i|$$

Example 1.4 (Function Space C[a,b]). The Function space consists of all real valued continuous functions on the closed interval $[a, b]$, written as $C[a, b]$. Let f and g be two continuous functions in $C[a, b]$. Then the distance between f and g is defined as:

$$d(f, g) = \max_{x \in [a,b]} |f(x) - g(x)|$$

Example 1.5 (Discrete Metric Space)**.** Let X be a set and d be a metric on X such that:

$$d(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

(X, d) is called a discrete metric space.

Example 1.6 (Space $B(A)$ of bounded real or complex valued functions on a set A)**.** The metric space consists of all bounded real or complex valued functions on a set A, written as $B(A)$. Let $x : A \to \mathbb{R}$ and $y : A \to \mathbb{R}$ be two bounded functions in $B(A)$. Then the distance between x and y is defined as:

$$d(x,y) = \sup_{a \in A} |x(a) - y(a)|$$

§1.2 ℓ^p space

Example 1.7 (Space ℓ^p)**.** Let $p \geq 1$ be a fixed real number. Let $x \in \ell^p$. Then $x = x_1, x_2, \ldots$ is a sequence of real numbers such that:

$$\sum_{i=1}^{\infty} |x_i|^p < \infty$$

Then the distance between x and y is defined as:

$$d(x,y) = \left(\sum_{i=1}^{\infty} |x_i - y_i|^p \right)^{\frac{1}{p}}$$

In the special case where $p = 2$, we have the hilbert sequence space l^2. In general ℓ^p is a metric space. We will prove prove this using Hölder's and Minkowski's inequality.

Lemma 1.1 (Hölder's Inequality)

Let $x = \{x_i\} \in \ell^p$ and $y = \{y_i\} \in \ell^q$ be two sequences of complex numbers where $p > 1$ is a real number and q be the conjugate exponent of p i.e.

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{1.1}$$

Then the Hölder's inequality states that:

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q \right)^{\frac{1}{q}} \tag{1.2}$$

Proof. From equation (1.1), we have

$$\begin{aligned} 1 &= \frac{p+q}{pq} \\ &\implies pq = p + q \\ &\implies -p - q + pq + 1 = 1 \\ &\implies p(q-1) - 1(q-1) = 1 \\ &\implies (p-1)(q-1) = 1 && (1.3a) \\ &\implies \frac{1}{p-1} = q - 1 && (1.3b) \end{aligned}$$

Let $u = t^{p-1}$, then using equation (1.3b) we have

$$\begin{aligned} u &= t^{p-1} \\ \implies t &= u^{\frac{1}{p-1}} \\ \implies t &= u^{q-1} && (1.4) \end{aligned}$$

Let α, β be any positive real numbers. Then $\alpha\beta$ is the area of the rectangle in the figures below.

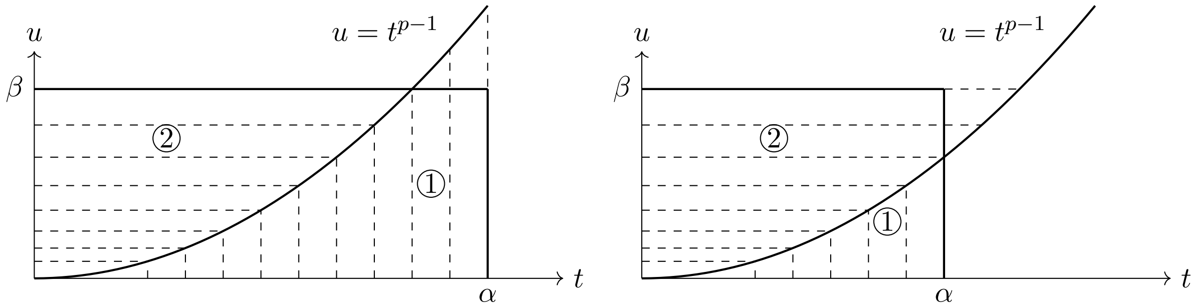


Figure 1.1: In both figures, Area (1): $\int_0^\alpha t^{p-1}dt$, Area(2): $\int_0^\beta u^{q-1}dt$ (Indicated by dashed lines)

In both figures of (1.1),

$$\alpha\beta \leq \int_0^\alpha t^{p-1}dt + \int_0^\beta q^{q-1}dt = \frac{\alpha^p}{p} + \frac{\beta^q}{q} \tag{1.5}$$

Note that, this holds trivially for $\alpha = 0$ or $\beta = 0$.

Let $\tilde{x} = \{\tilde{x}_i\}$ and $\tilde{y} = \{\tilde{y}_i\}$ be two sequences of complex numbers such that,

$$\sum_{i=1}^{\infty} |\tilde{x}_i|^p = 1, \quad \sum_{i=1}^{\infty} |\tilde{y}_i|^q = 1 \tag{1.6}$$

Setting $\alpha = |\tilde{x}_i|, \beta = |\tilde{y}_i|$ for a given i, the inequality (1.5) becomes,

$$\begin{aligned} |\tilde{x}_i\tilde{y}_i| &\leq \frac{|\tilde{x}_i|^p}{p} + \frac{|\tilde{y}_i|^q}{q} \\ \implies |\tilde{x}_i\tilde{y}_i| &\leq \frac{1}{p}|\tilde{x}_i|^p + \frac{1}{q}|\tilde{y}_i|^q \end{aligned} \tag{1.7}$$

Summing over all i from 1 to ∞ in equation (1.7), and using equation (1.6) and (1.1), we get,

$$\begin{aligned} \sum_{i=1}^{\infty} |\tilde{x}_i\tilde{y}_i| &\leq \sum_{i=1}^{\infty} \left(\frac{1}{p}|\tilde{x}_i|^p + \frac{1}{q}|\tilde{y}_i|^q\right) \\ \implies \sum_{i=1}^{\infty} |\tilde{x}_i\tilde{y}_i| &\leq \frac{1}{p}\sum_{i=1}^{\infty} |\tilde{x}_i|^p + \frac{1}{q}\sum_{i=1}^{\infty} |\tilde{y}_i|^q \\ \implies \sum_{i=1}^{\infty} |\tilde{x}_i\tilde{y}_i| &\leq \frac{1}{p} + \frac{1}{q} = 1 \end{aligned} \tag{1.8}$$

Note that, we were able to take the sum to infinity since the series on the right hand side converges. Hence, the series on the left converges by comparison test.

We now take nonzero $x = \{x_i\}$ and $y = \{y_i\}$ in ℓ^p and set,

$$\tilde{x}_i = \frac{x_i}{\left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}}}, \quad \tilde{y}_i = \frac{y_i}{\left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}} \tag{1.9}$$

It is immediate that (1.9) satisfies the conditions of (1.6). Substituting (1.9) in (1.8), we get,

$$\begin{aligned} &\frac{\sum_{i=1}^{\infty} |x_iy_i|}{\left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}} \leq 1 \\ \implies &\sum_{i=1}^{\infty} |x_iy_i| \leq \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}} \end{aligned}$$

■

Corollary 1.2 (Cauchy-Schawrz Inequality)
Let $x = \{x_i\} \in l^2$ and $y = \{y_i\} \in l^2$ be two sequences of complex numbers. Then the Cauchy-Schawrz inequality states that:

$$\left(\sum_{i=1}^{\infty} |x_i y_i|\right)^2 \leq \sqrt{\sum_{j=1}^{\infty} |x_j|^2} \sqrt{\sum_{k=1}^{\infty} |y_k|^2} \tag{1.10}$$

Proof. Setting $p = q = 2$ in the Hölder's inequality (1.2), we get the Cauchy-Schawrz inequality. ■

Lemma 1.3 (Minkowski's Inequality)
Let $x = \{x_i\} \in \ell^p$ and $y = \{y_i\} \in \ell^p$ be two sequences of complex numbers. Let $p \geq 1$ be a real number. Then the Minkowski's inequality states that:

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}} \tag{1.11}$$

Proof. Let $p = 1$, then from triangle inequality we have,

$$|x_i + y_i| \leq |x_i| + |y_i| \tag{1.12}$$

summing over all i from 1 to a fixed n, we get,

$$\sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i|$$

If we take the limit as $n \to \infty$, both series on the right converges since $x, y \in \ell^p$. Hence, by comparison test, infinite series on the left also converges and we get,

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|\right) \leq \left(\sum_{i=1}^{\infty} |x_i|\right) + \left(\sum_{i=1}^{\infty} |y_i|\right)$$

Which is the Minkowski's inequality for $p = 1$.

Let $p > 1$, and let $z_i = x_i + y_i$ for all i. If z_i is the zero sequence, then the inequality is trivially true. Hence, we assume that z_i is not the zero sequence. Then from triangle inequality we have,

$$\begin{aligned} |z_i|^p &= |z_i| \cdot |z_i|^{p-1} \\ &= |x_i + y_i| \cdot |z_i|^{p-1} \\ &\leq (|x_i| + |y_i|) \cdot |z_i|^{p-1} \text{ (Using equation (1.12))} \\ &= |x_i| \cdot |z_i|^{p-1} + |y_i| \cdot |z_i|^{p-1} \end{aligned} \tag{1.13}$$

Summing over all i from 1 to a fixed n, we get,

$$\sum_{i=1}^{n} |z_i|^p \leq \sum_{i=1}^{n} |x_i| \cdot |z_i|^{p-1} + \sum_{i=1}^{n} |y_i| \cdot |z_i|^{p-1} \tag{1.14}$$

Consider the finite sequnces $(|x_i|)_{i=1}^n$, which is an ℓ^p sequence and $(|z_i|^{p-1})_{i=1}^n$, which is an ℓ^q sequence with q being the conjugate exponent of p. Hence,

$$\begin{aligned} &\frac{1}{p} + \frac{1}{q} = 1 \\ \implies &pq = p + q \\ \implies &pq - q = p \\ \implies &(p-1)q = p \end{aligned} \tag{1.15}$$

Using Hölder's inequality (1.2) we get,

$$\sum_{i=1}^{n} ||x_i| \cdot |z_i|^{p-1}| \leq \left(\sum_{j=1}^{n} ||x_j||^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} ||z_k|^{(p-1)}|^q\right)^{\frac{1}{q}}$$
$$\implies \sum_{i=1}^{n} |x_i| \cdot |z_i|^{p-1} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |z_k|^{(p-1)q}\right)^{\frac{1}{q}}$$
$$\implies \sum_{i=1}^{n} |x_i| \cdot |z_i|^{p-1} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |z_k|^{p}\right)^{\frac{1}{q}} \tag{1.16}$$

Similarly, for finite sequence $(|y_i|)_{i=1}^{n}$ and $(|z_i|^{p-1})_{i=1}^{n}$, we get,

$$\sum_{i=1}^{n} |y_i| \cdot |z_i|^{p-1} \leq \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |z_k|^{p}\right)^{\frac{1}{q}} \tag{1.17}$$

Therefore, from equations (1.14), (1.16) and (1.17), we get,

$$\sum_{i=1}^{n} |z_i|^p \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |z_k|^{p}\right)^{\frac{1}{q}} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |z_k|^{p}\right)^{\frac{1}{q}}$$
$$\implies \sum_{i=1}^{n} |z_i|^p \leq \left(\left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}}\right) \cdot \left(\sum_{k=1}^{n} |z_k|^{p}\right)^{\frac{1}{q}} \tag{1.18}$$

Since, z_k is not the zero sequence, $(\sum_{k=1}^{n} |z_k|^p)^{\frac{1}{q}}$ is nonzero. Therefore we can divide both side by $(\sum_{k=1}^{n} |z_k|^p)^{\frac{1}{q}}$ and get,

$$\frac{\sum_{i=1}^{n} |z_i|^p}{(\sum_{k=1}^{n} |z_k|^p)^{\frac{1}{q}}} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}}$$
$$\implies \frac{\sum_{i=1}^{n} |z_i|^p}{(\sum_{i=1}^{n} |z_i|^p)^{\frac{1}{q}}} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}}$$
$$\implies \left(\sum_{i=1}^{n} |z_i|^p\right)^{1-\frac{1}{q}} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}}$$
$$\implies \left(\sum_{i=1}^{n} |z_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}} \quad \left[\text{Since } \frac{1}{p} + \frac{1}{q} = 1\right] \tag{1.19}$$

If we taking the limit as $n \to \infty$, both series on the right converges since $x, y \in \ell^p$. Hence, by comparison test, infinite series on the left also converges and we get,

$$\left(\sum_{i=1}^{\infty} |z_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{\infty} |y_j|^p\right)^{\frac{1}{p}}$$
$$\implies \left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{\infty} |y_j|^p\right)^{\frac{1}{p}}$$
$$\implies \left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}}$$

which is the Minkowski's inequality for $p > 1$. ■

Now we will prove that ℓ^p is a metric space with the metric defined in example (1.7).

Proposition 1.4 (ℓ^p is a metric space)
Let $p \geq 1$ be a fixed real number. Then the sequence space ℓ^p is a metric space with the metric define as,

$$d(x, y) = \left(\sum_{i=1}^{\infty} |x_i - y_i|^p \right)^{\frac{1}{p}} \tag{1.20}$$

Proof. We need to show that d satisfies all four properties in definition (1.1).

Since, $\sum_{i=1}^{\infty} |x_i - y_i|^p$ is a series of absolute values, it is non-negative. Hence, $d(x, y) \geq 0$.

Since $y \in \ell^p$, $-y = \{-y_i\} \in \ell^p$. Hence, using Minkowski's inequality for sequences x and $-y$, we get,

$$\begin{aligned} \left(\sum_{i=1}^{\infty} |x_i + (-y_i)|^p \right)^{\frac{1}{p}} &\leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |-y_i|^p \right)^{\frac{1}{p}} \\ \implies \left(\sum_{i=1}^{\infty} |x_i - y_i|^p \right)^{\frac{1}{p}} &\leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |y_i|^p \right)^{\frac{1}{p}} \end{aligned}$$

The series on the right converges since $x, y \in \ell^p$. Hence, the series on the left also converges by comparison test. Therefore $0 \leq d(x, y) < \infty$, hence d satisfies property (M1).

Property (M2) and (M3) are trivially satisfied by the definition of d. We need to prove property (M4), i.e. the triangle inequality. Given $x, y, z \in \ell^p$, we have,

$$\begin{aligned} d(x, y) &= \left(\sum_{i=1}^{\infty} |x_i - y_i|^p \right)^{\frac{1}{p}} \\ &\leq \left(\sum_{i=1}^{\infty} (|x_i - z_i| + |z_i - y_i|)^p \right)^{\frac{1}{p}} \\ &\leq \left(\sum_{i=1}^{\infty} |x_i - z_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |z_i - y_i|^p \right)^{\frac{1}{p}} \qquad \text{[Using Minkowski's inequality]} \\ &= d(x, z) + d(z, y) \end{aligned}$$

Hence, d satisfies property (M4). Therefore, ℓ^p is a metric space. ■

2 Normed Space and Banach Space

§2.1 Normed Space, Banach Space

Definition 2.1 (Normed Space, Banach Space)**.** A **Normed space** $(X, ||\cdot||)$ is a vector space X with a norm $||\cdot||$ defined on it. A **Banach space** is a complete metric space. A **norm** on a real or a complex vector space$X$ is a real valued function on X whose value at $x \in X$ is denotes by

$$||x||$$

and which has the flowing properties,

- **(N1)** $||x|| \geq 0$
- **(N2)** $||x|| = 0 \implies x = 0$
- **(N3)** $||\alpha x|| = \alpha ||x||$
- **(N4)** $||x + y|| \leq ||x|| + ||y||$ (Triangle inequality)

Here X and y are arbitrary vectors in X and α is any scalar.

A norm on $(X, ||\cdot||)$ defines a metric on (X, d) which is defined by

$$d(x, y) = ||x - y|| \tag{2.1}$$

and is called **The metric induced by the norm**. Thus the normed space $(X, ||\cdot||)$ is a topological space w.r.t the metric induced from the norm. For later use, we note that

$$||x|| = ||x - y + y|| \leq ||x - y|| + ||y|| \quad \text{by } \textbf{(N4)}$$

So that

$$||x|| - ||y|| \leq ||x - y|| \tag{2.2}$$

also

$$||y|| = ||y - x + x|| \leq ||y - x|| + ||x|| = ||x - y|| + ||x|| \quad \text{again by } \textbf{(N4)}$$

Then,

$$-(||x|| - ||y||) \leq ||x - y|| \implies ||x|| - ||y|| \geq -||x - y|| \tag{2.3}$$

(2.2) and (2.3) together imply

$$-||x - y|| \leq ||x|| - ||y|| \leq -||x - y|| \tag{2.4}$$

equation (2.4) implies that

$$|||x|| - ||y||| \leq ||x - y|| \tag{2.5}$$

Inequality (2.4) can be used to prove that the norm function

$$||\cdot|| : (X, ||\cdot||) \to \mathbb{R}$$

i.e., $x \mapsto ||x||$ is a continuous mapping.

§2.1.i Examples of normed spaces

Example 2.1. Euclidean space $\mathbb{R}$ and the unitary space $\mathbb{C}$ are normed spaces. In fact, they are normed space defined by

$$||x|| = \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{\frac{1}{2}} = \sqrt{|\xi_1|^2 + ... + |\xi_n|^2} \tag{2.6}$$

Both of them are complete in the metric (2.1) defined by the norm given in (2.6):

$$d(x, y) = ||x - y|| = \sqrt{|\xi_1 - \eta_1|^2 + ... + |\xi_n - \eta_n|^2} \tag{2.7}$$

there, $x = (\xi_1, ..., \xi_n)$ and $y = (\eta_1, ..., \eta_n)$ are vectors of $\mathbb{R}$ or $\mathbb{C}$, i.e., they are real or complex n-tuples depending on if they are element of $\mathbb{R}^n$ or $\mathbb{C}^n$, respectively. Note that here $\mathbb{R}^n$ or $\mathbb{C}^n$ are treated as vector spaces over the field of real numbers or complex numbers, respectively in contract to bare sets $(\mathbb{R}^n, d_{\text{Euc}})$ or $(\mathbb{C}^n, d_{\text{uni}})$. To reduce confusion, one can also denote the Banach space by $(\mathbb{R}^n, ||\cdot||_{\text{Euc}})$ to differentiate it from the metric space $(\mathbb{R}^n, ||\cdot||_{\text{Euc}})$. In the former case $\mathbb{R}^n$ is a real vector space while in the latter case we require $\mathbb{R}^n$ to be a bare set of n-tuples by real numbers.

Example 2.2. (Normed space ℓ^p) The set ℓ^p was introduced in the 5^{th} example of the first lecture. It is the set of p-summable sequences of complex numbers, i.e., if $x = (\xi_i) \in \ell^p$, then

$$\sum_{j=1}^{\infty} |\xi_j|^p < \infty \tag{2.8}$$

ℓ^p has a complex vector space structure. Vector addition and scalar multiplication on ℓ^p is given by:

$$\begin{gathered}(\xi_1, \xi_2, ..) + (\eta_1, \eta_2, ...) = (\xi_1 + \eta_1, \xi_2 + \eta_2, ...)\\ \alpha(\xi_1, \xi_2, ...) = (\alpha\xi_1, \alpha\xi_2, ...)\end{gathered} \tag{2.9}$$

where $(\xi_j), (\eta_j) \in \ell^p$. You should verify that all the axioms of a vector space are fulfilled under definition (2.9). It is actually a Banach space with the norm given by

$$||x|| = \left(\sum_{j=1}^{\infty} |\xi_j|^p\right)^{\frac{1}{p}} \tag{2.10}$$

where $x = (\xi_j) \in \ell^p$. This norm induces the metric

$$d(x, y) = ||x - y|| = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p\right)^{\frac{1}{p}} \tag{2.11}$$

With $x = (\xi_j)$ and $y = (\eta_j)$ are both elements of ℓ^p. This metric given by equation (2.9) of the first lecture while discussing the metric space structure of ℓ^p. Here, we see that it is not just a bare set. It has the structure of a complex vector space and the norm given by (2.10) induces the already known metric space structure on it. And we know that it is complete with respect to that already known metric which indicates that ℓ^p is a Banach space.

Example 2.3. (Normed space ℓ^∞) The set has been introduced in example 2 of chapter 1. It is the set of all bounded sequences of complex numbers, i.e., every element of it is a complex sequence$x = (\xi_1, \xi_2, \dots)$, briefly $x = (\xi_j)$, such that for all $j = 1, 2, \dots$, one has

$$|\xi_j| \le C_x, \tag{2.12}$$

where C_x is a positive real number. The set ℓ^∞ can be endowed with the structure of a complex vector space by introducing the vector addition and scalar multiplication using (2.9). Then one defines a norm on the vector space ℓ^∞ by

$$||x|| = \sup_{j\in\mathbb{N}} |\xi_j| \tag{2.13}$$

for $x = (\xi_j) \in \ell^\infty$.

The norm given by (2.13) induces the metric

$$d(x,y) = \sup_{j\in\mathbb{N}} |\xi_j - \eta_j| \tag{2.14}$$

for $x = (\xi_j) \in \ell^\infty$ and $y = (\eta_j) \in \ell^\infty$ as introduced in example 2 of lecture 1 while discussing the metric space structure of ℓ^∞. And of course ℓ^∞ is complete with respect to the induced metric (2.15). In other words, ℓ^∞ is a **Banach space**.

§2.1.ii An Incomplete metric space and it's completion

The set of all complex-valued continuous functions on the closed interval $[a,b]$ can easily be endowed with the structure of a complex vector space that we also denote by $\mathcal{C}[a,b]$. The vector addition and scalar multiplication on $\mathcal{C}[a,b]$ is pointwise:

$$\begin{aligned}(x+y)(t) &= x(t) + y(t)\\ (\alpha x)(t) &= \alpha \cdot x(t)\end{aligned} \tag{2.15}$$

where $x, y \in \mathcal{C}[a,b]$, i.e., $x, y : [a,b] \to \mathbb{C}$ and α is a complex number. One can define a norm on $\mathcal{C}[a,b]$ by

$$\|x\|_p = \left(\int_a^b |x(t)|^p \, dt\right)^{\frac{1}{p}} \tag{2.16}$$

The normed space $(\mathcal{C}[a,b], \|\cdot\|_p)$ is not complete, as was seen in example 8 of lecture 1. The space $(\mathcal{C}[a,b], \|\cdot\|_p)$ can be completed by Theorem 2.6.

The completed metric space is actually a normed space that will follow from a more general completion theorem for normed spaces (to be discussed in a while). The norm completion of $(\mathcal{C}[a,b], \|\cdot\|_p)$ is denoted by $L^p[a,b]$ and is called the normed space of p-integrable functions. The case $p = 2$ is even more interesting and $\ell^2[a,b]$ has the structure of something called a Hilbert space to be discussed in great detail later.

§2.1.iii Example of metric space (X,d) whose metric can not be induced from a norm

Suppose X is any vector space. Put the discrete metric d_{dis} on it so that (X, d_{dis}) is a metric space. We show that there exists no norm on the vector space X that will induce the discrete metric on it in the sense of equation (2.1). Suppose the contrary, i.e., there exists a norm denoted by $\|\cdot\|_{\text{dis}}$ that will induce d_{dis} on X:

$$d_{\text{dis}}(x,y) = \|x-y\|_{\text{dis}} \tag{2.17}$$

for all $x, y \in X$. Choose $x, y \in X$ such that $x \neq y$ and α is a nonzero scalar. Then $\alpha x \neq \alpha y$ so that

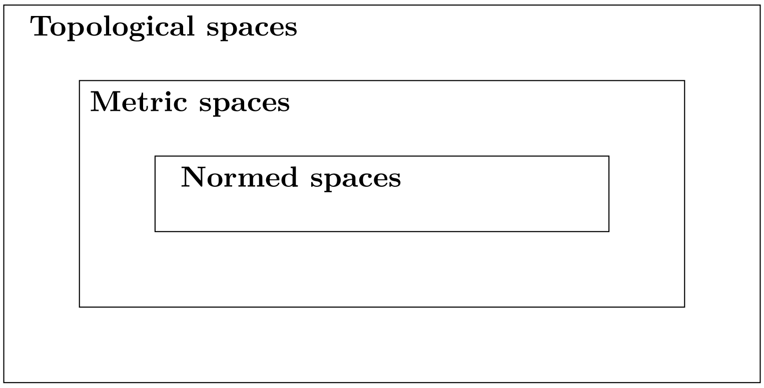
$$d_{\text{dis}}(\alpha x, \alpha y) = 1, \quad \text{[from the definition of discrete metric]} \tag{2.18}$$

But

$$\begin{aligned} d_{\text{dis}}(\alpha x, \alpha y) &= \|\alpha x - \alpha y\|_{\text{dis}} = \|\alpha(x-y)\|_{\text{dis}} && (2.19)\\ &= |\alpha|\|x-y\|_{\text{dis}} \quad \text{[From (N3)]} && (2.20)\\ &= |\alpha| d_{\text{dis}}(x,y) && (2.21)\\ &= |\alpha| && (2.22)\end{aligned}$$

(2.18) and (2.19) are in contradiction with each other as $|\alpha| \neq 1$, by hypothesis. Hence, there exists no norm on the vector space X that will induce d_{dis} on X.

In the light of the discrete metric topology on any vector space, we thus see that there are metric spaces that are not normed spaces. So far, we have found the following hierarchy:



Lemma 2.1

(Translational invariance) A metric d induced by a norm on a normed space $(X, \|\cdot\|)$ satisfies

$$\begin{aligned} &\text{(a)} \quad d(x+a, y+a) = d(x, y), \\ &\text{(b)} \quad d(\alpha x, \alpha y) = |\alpha| d(x, y), \end{aligned}$$

for all $x, y, a \in X$ and for all scalars α.

Proof.

$$\begin{aligned} d(x+a, y+a) &= \|(x+a) - (y+a)\| && \text{By eq. (2.1)} \\ &= \|x - y\| = d(x, y). \end{aligned}$$

And

$$\begin{aligned} d(\alpha x, \alpha y) &= \|\alpha x - \alpha y\| && \text{again by eq. (2.1)} \\ &= |\alpha| \|x - y\| && \text{by (N3) property of norm} \\ &= |\alpha| d(x, y). \end{aligned}$$

∎

§2.2 Finite dimensional normed spaces and their subspaces

Lemma 2.2 (Linear combinations)

Let $\{x_1, \ldots, x_n\}$ be a linearly independent set of vectors in a normed space $(X, \|\cdot\|)$. Then there exists a real number $c > 0$ such that for every choice of scalars $\alpha_1, \ldots, \alpha_n$, one has

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|) \tag{2.23}$$

the left side is the lower bound of this norm.

Theorem 2.3 (Completeness)

Every finite dimensional subspace $(Y, \|\cdot\|_Y)$ of a normed space $(X, \|\cdot\|_X)$ is complete. In particular, every finite dimensional normed space is complete.

Theorem 2.4 (Closedness)
Every finite dimensional subspace $(Y, \|\cdot\|_Y)$ of a normed space $(X, \|\cdot\|_X)$ is closed in $(X, \|\cdot\|_X)$.

Example 2.4. Consider the vector space of all complex-valued continuous functions defined on the closed interval $[0,1]$ of $\mathbb{R}$ and equip it with the sup norm:

$$\|x\|_\infty = \sup_{t\in[0,1]} |x(t)| \tag{2.24}$$

Then $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ is a Banach space. Let $Y = \operatorname{span}\{x_0, x_1, x_2, \ldots\}$ where $x_j(t) = t^j$, so that Y is the set of all polynomials. Then $(Y, \|\cdot\|_\infty|_Y)$ is not closed in $(\mathcal{C}[0,1], \|\cdot\|_\infty)$.

§2.3 Further Properties of a normed space

By a subspace of a normed space $(X, \|\cdot\|_X)$, we mean a vector subspace Y of the vector space X, with the norm obtained by restricting the norm $\|\cdot\|_X$ to the subset Y. Let us denote this restriction by $\|\cdot\|_{X|_Y}$, and the corresponding normed subspace by the pair $(Y, \|\cdot\|_{X|_Y})$. The vector space X is equipped with the metric topology with the metric induced from the norm $\|\cdot\|_X$. When Y is closed in X with respect to the above-mentioned metric topology, we say that $(Y, \|\cdot\|_{X|_Y})$ is a *closed subspace* of the normed space $(X, \|\cdot\|_X)$.

Theorem 2.5
(Subspace of a Banach space) A subspace $(Y, \|\cdot\|_{X_Y})$ of a Banach space $(X, \|\cdot\|_X)$ is complete if and only if $(Y, \|\cdot\|_{X_Y})$ is a closed subspace of the normed space $(X, \|\cdot\|_X)$.

Proof. This proof actually follows from analogous results on metric space. ■

§2.3.i Convergence of sequences in normed spaces

Definition 2.2. A sequence (x_n) in a normed space $(X, \|\cdot\|)$ is **convergent** if X contains an x such that

$$\lim_{n\to\infty} \|x_n - x\| = \lim_{n\to\infty} d_X(x_n, x) = 0 \tag{2.25}$$

In (2.25), d_X is the metric induced from the norm $\|\cdot\|$ of the normed space $(X, \|\cdot\|)$. One then writes $x_n \to x$ and calls x the **limit** of (x_n).

Definition 2.3. A sequence (x_n) in a normed space $(X, \|\cdot\|)$ is *Cauchy* if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\|x_n - x_m\| = d_X(x_n, x_m) < \epsilon, \quad \forall m, n \geq N \tag{2.26}$$

Sequences were available to us even in a *general metric space*. In a normed space, we have the additional notion of *series*.

If (x_k) is a sequence in a normed space $(X, \|\cdot\|)$, then each x_k is a vector in X, and we can associate with (x_k) the sequence (s_n) of *partial sums*:

$$s_n = x_1 + x_2 + \cdots + x_n \tag{2.27}$$

Here, $+$ denotes vector addition in $(X, \|\cdot\|)$.

If (s_n) is convergent, say

$$S_n \to s, \text{ that is, } \|s_n - s\| \to 0,$$

then the *infinite series*, or briefly, *series*

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \dots \tag{2.28}$$

is said to *converge* or to be *convergent*; s is called the *sum* of the series, and we write

$$s = \sum_{k=1}^{\infty} x_k = x_1 + x_2 + \dots \tag{2.29}$$

If $\|x_1\| + \|x_2\| + \dots$ converges, the series (2.28) is said to be *absolutely convergent*.
The concept of convergence of series in a normed space $(X, \|\cdot\|)$ helps us define a basis as follows.

Definition 2.4. If a normed space $(X, \|\cdot\|)$ contains a sequence (e_n) with the property that for every $x \in X$, there exists a sequence of scalars (α_n) such that

$$\lim_{n\to\infty} \|x - (\alpha_1 e_1 + \dots + \alpha_n e_n)\| = 0 \tag{2.30}$$

then (e_n) is called a *Schauder basis* for $(X, \|\cdot\|)$.

The series $\sum_{k=1}^{\infty} \alpha_k e_k$, which has the sum x, is then called the **expansion** of x with respect to (e_n), and we write

$$x = \sum_{k=1}^{\infty} \alpha_k e_k \tag{2.31}$$

For example, ℓ^p has a Schauder basis, namely (e_n), where

$$e_n = (\delta_{nj}),$$

i.e., e_n is the sequence whose n-th term is 1 and all other terms are zero. Thus,

$$\begin{aligned} e_1 &= (1, 0, 0, 0, \dots), \\ e_2 &= (0, 1, 0, 0, \dots), \\ e_3 &= (0, 0, 1, 0, \dots), \end{aligned}$$

etc.

Theorem 2.6 (Completion)
Let $(X, \|\cdot\|_X)$ be a normed space. Then there is a Banach space $(\hat{X}, \|\cdot\|_{\hat{X}})$ and an isometry $J : (X, \|\cdot\|_X) \to (W, \|\cdot\|_{\hat{X}})$ of $(X, \|\cdot\|_X)$ to a dense subspace $(W, \|\cdot\|_{\hat{X}})$ of $(\hat{X}, \|\cdot\|_{\hat{X}})$. The space $(\hat{X}, \|\cdot\|_{\hat{X}})$ is unique up to isometries.

§2.4 Finite dimensional normed spaces and their subspaces

Lemma 2.7 (Linear combinations)
Let $\{x_1, \dots, x_n\}$ be a linearly independent set of vectors in a normed space $(X, \|\cdot\|)$. Then there exists a real number $c > 0$ such that for every choice of scalars $\alpha_1, \dots, \alpha_n$, one has

$$\|\alpha_1 x_1 + \dots + \alpha_n x_n\| \geq c(|\alpha_1| + \dots + |\alpha_n|), \tag{2.32}$$

i.e. the right side is a lower bound of this norm.

Theorem 2.8 (Completeness)
Every finite dimensional subspace $(Y, \|\cdot\|_Y)$ of a normed space $(X, \|\cdot\|_X)$ is complete. In particular, every finite dimensional normed space is complete.

Theorem 2.9 (Closedness)
Every finite dimensional subspace $(Y, \|\cdot\|_Y)$ of a normed space $(X, \|\cdot\|_X)$ is closed in $(X, \|\cdot\|_X)$.

Example 2.5. Consider the vector space of all complex-valued continuous functions defined on the closed interval $[0, 1]$ of $\mathbb{R}$ and equip it with the sup norm:

$$\|x\|_\infty = \sup_{t\in[0,1]} |x(t)| \tag{2.33}$$

Then $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$ is a Banach space. Let $Y = \operatorname{span}\{x_0, x_1, x_2, \dots\}$ where $x_j(t) = t^j$, so that Y is the set of all polynomials. Then $(Y, \|\cdot\|_\infty|_Y)$ is not closed in $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$.

§2.5 Compactness and Finite Dimensions

Definition 2.5. A metric space (X, d_X) is said to be *compact* if every sequence in (X, d_X) has a convergent subsequence.

Lemma 2.10
A compact subset M of a metric space (X, d_X) is closed and bounded.

The converse of Lemma 2.10 is not true. Consider the metric space ℓ^2. Now, consider the sequence (e_n) in ℓ^2, where $e_n = (\delta_{nj}) \in \ell^2$, i.e., the n-th term e_n of (e_n) is itself a sequence that has 1 in the n-th term and 0 elsewhere. The sequence (e_n) is bounded. We know that ℓ^2 is a normed space so that the metric here is induced from the norm $\|\cdot\|_2$ given by

$$\|x\| = \left(\sum_{k=1}^{\infty} |x_k|^2\right)^{1/2} \tag{2.34}$$

with $x = (x_k) \in \ell^2$. And $\|e_n\| = 1$, for each n. Also, the point set for the sequence (e_n) has its diameter bounded by 1, and hence (e_n) is a bounded sequence in ℓ^2.

Distinct points of the sequence are separated by a distance of 1, i.e., if $m \neq n$, then

$$d(e_m, e_n) = 1,$$

so that every point of (e_n) has an ϵ-neighborhood that contains no other points of (e_n) except for the point itself, meaning that none of the points of (e_n) is an accumulation point so that the point set of (e_n) is closed and bounded in ℓ^2. But it is not compact in ℓ^2, as (e_n) itself, being a sequence in the point set of (e_n), does not have a convergent subsequence.

For a finite dimensional normed space, we have the following nice result:

Theorem 2.11
In a finite dimensional normed space $(X, \|\cdot\|_X)$, any subset $M \subset X$ is compact if and only if it is closed and bounded.

Another important result is due to Riesz:

Lemma 2.12 (Riesz's Lemma)
Let $(y, \|\cdot\|_y)$ and $(z, \|\cdot\|_z)$ be subspaces of a normed space $(X, \|\cdot\|_X)$ of any dimension. And suppose that $Y \subset Z$ is closed. Then for every real number $\theta \in (0,1)$, there exists $z \in Z$ such that

$$\|z\| = 1, \quad \|z - y\| \geq \theta \quad \forall y \in Y \tag{2.35}$$

Proof. Consider $v \in Z \setminus Y$ and denote its distance from Y by a,

$$a = \inf_{y \in Y} \|v - y\| \tag{2.36}$$

Since Y is closed, v can't be in $\overline{Y}$ as $v \in Z \setminus Y$. Hence, $a > 0$.
We now choose $\theta \in (0,1)$. By the definition of an infimum, $\exists\, y_0 \in Y$ such that

$$a < \|v - y_0\| \leq \frac{a}{\theta} \quad (\text{note that } 0 < \theta < 1, \text{ one has } \frac{a}{\theta} > a) \tag{2.37}$$

$$\text{Let } z = c(v - y_0) \text{ with } c = \frac{1}{\|v - y_0\|}, \tag{2.38}$$

so that $\|z\| = 1$, and we show that

$$\|z - y\| = \|c(v - y_0) - y\| \tag{2.39}$$
$$= c\|v - y_0 - y\| \geq c \cdot a \tag{2.40}$$
$$= \frac{a}{\|v - y_0\|} \tag{2.41}$$

Now, since y_0, $y \in Y$, $y_0 + \frac{y}{2} \in Y$, i.e., $y_1 \in Y$. Hence, by the definition (2.36) of a,

$$a \leq \|v - y_1\| \tag{2.42}$$

Therefore, from (2.39), one obtains

$$\|z - y\| = c\|v - y_0\| \geq a = \frac{a}{\|v - y_0\|} \tag{2.43}$$

Now, (2.37) yields,

$$\|v - y_0\| \leq \frac{a}{\theta} \tag{2.44}$$
$$\implies \frac{1}{\|v - y_0\|} \geq \frac{1}{\frac{a}{\theta}} \tag{2.45}$$
$$\implies \frac{a}{\|v - y_0\|} \geq \frac{a}{\frac{a}{\theta}} = \theta \tag{2.46}$$

Now, combining (2.43) with (2.44), one obtains,

$$\|z - y\| > 0.$$

Since $y \in Y$ was arbitrary, it completes the proof. ■

Now, if one has a closed unit ball in a finite dimensional normed space, then the closed unit ball, being closed and bounded, is also compact by Theorem 2.11. The converse of this statement can be proved using Riesz's Lemma.

Theorem 2.13
If a normed space $(X, \|\cdot\|_X)$ has the property that the closed unit ball $M = \{z \in X \mid \|z\|_X \leq 1\}$ is compact, then X is finite dimensional.

Proof. We assume that M is compact but X is infinite dimensional. Our goal would be to reach a contradiction. First, take x_1 from X of norm 1. This x_1 then generates a 1-dimensional subspace $(X_1, \|\cdot\|_{X_1})$ where $X_1 = \{cx_1 \mid c \text{ is any scalar}\}$. Now, by Theorem 2.9, X_1, being finite dimensional, is closed in $(X, \|\cdot\|_X)$. Since X is infinite dimensional, the containment $X_1 \subset X$ is proper. Then, by Riesz's Lemma, $\exists x_2 \in X$ of norm 1 such that

$$\|x_2 - x_1\|_X \geq \frac{1}{2} \tag{2.47}$$

Now, x_1, x_2 generate a 2-dimensional proper closed subspace $(X_2, \|\cdot\|_{X_2})$ of $(X, \|\cdot\|_X)$, i.e.,

$$X_2 = \{c_1x_1 + c_2x_2 \in X \mid c_1, c_2 \text{ are any scalars}\} \tag{2.48}$$

Then, by Riesz's Lemma, there is $x_3 \in X$ with $\|x_3\| = 1$ such that $\forall z \in X_2$, one has

$$\|x_3 - z\|_X \geq \frac{1}{2} \tag{2.49}$$

In particular,

$$\|x_3 - x_1\|_X \geq \frac{1}{2} \text{ and } \|x_3 - x_2\|_X \geq \frac{1}{2} \tag{2.50}$$

Now, proceeding by induction, we obtain a sequence (x_n) of elements $x_n \in M$ s.t.

$$\|x_m - x_n\| \geq \frac{1}{2} \quad \text{whenever } m \neq n.$$

Obviously (x_n) can't have a convergent subsequence. This contradicts the compactness of M. Hence, our assumption $\dim X = \infty$ was false, and $\dim X < \infty$ ■

Remark 2.1. Theorem 2.13 has various applications. We'll use it as a basic tool in the study of compact operators.

Theorem 2.14

Let $T : (X, d_X) \to (Y, d_Y)$ be a continuous mapping between metric spaces. Then the image of a compact subset M of X under T is compact.

Corollary 2.15

(Maximum and minimum) A continuous mapping T of compact subset M of a metric space (X, d_X) to $\mathbb{R}$ assumes a maximum and a minimum at some points of M.

§2.6 Linear Operator

Definition 2.6 (Linear Operator)**.** A linear operator $T : \mathcal{D}(T) \to Y$ is a mapping between vector spaces where $\mathcal{D}(T) \subset X$ is a vector subspace. The range $\mathcal{R}(T)$ of T is contained in the vector space Y. All the vector spaces considered are defined over the same field. The mapping $T : \mathcal{D}(T) \to Y$ satisfies,

$$\begin{aligned} T(x+y) &= Tx + Ty \\ T(\alpha x) &= \alpha Tx \end{aligned} \tag{2.51}$$

for all $x, y \in \mathcal{D}(T)$ and scalars α.

Notation. $\mathcal{D}(T)$ denotes the domain of the linear operator T. The domain $\mathcal{R}(T)$ denotes the range. The null space or the kernal of T is denoted by $\mathcal{N}(T)$. That is, $\mathcal{N}(T) = \{x \in \mathcal{D}(T) : Tx = 0_Y\}$.

Theorem 2.16 (Range and Null Space)
If T is a linear operator then,

(a) The range $\mathcal{R}(T)$ is a vector space.

(b) If $\dim \mathcal{D}(T) = n < \infty$, then $\dim \mathcal{R}(T) \leq n$.

(c) The null space $\mathcal{N}(T)$ is a vector space.

Proof. (a) Let $y_1, y_2 \in \mathcal{R}(T)$. We want to show that $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$ for any scalars α, β. Since $y_1, y_2 \in \mathcal{R}(T)$, we have $y_1 = Tx_1, y_2 = Tx_2$ for some $x_1, x_2 \in \mathcal{D}(T)$. Since, $\mathcal{D}(T)$ is a vector space (by definition of linear operator), $\alpha x_1 + \beta x_2 \in \mathcal{D}(T)$ Then by linearity of T we have,

$$\begin{aligned} T(\alpha x_1 + \beta x_2) &= \alpha T x_1 + \beta T x_2 \\ &= \alpha y_1 + \beta y_2 \end{aligned}$$

Therefore, $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$. Hence, $\mathcal{R}(T)$ is a vector space.

(b) Choose $n+1$ arbitrary elements $y_1, y_2, \cdots y_{n+1} \in \mathcal{R}(T)$. Then, $y_1 = Tx_1, y_2 = Tx_2, \cdots y_{n+1} = Tx_{n+1}$ for some $x_1, x_2, \cdots x_{n+1} \in \mathcal{D}(T)$. Since dimension of $\mathcal{R}(T)$ is n, $\{x_1, x_2, \cdots x_{n+1}\}$ is linearly dependent. Hence,

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_{n+1} x_{n+1} = 0$$

for some scalars $\alpha_1, \alpha_2, \cdots \alpha_{n+1}$, not all zero. Then,

$$\begin{aligned} T(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_{n+1} x_{n+1}) &= \alpha T x_1 + \alpha T x_2 + \cdots + \alpha T x_{n+1} \\ &= \alpha y_1 + \alpha y_2 + \cdots + \alpha y_{n+1} = 0 \end{aligned}$$

Therefore, $\{y_1, y_2, \cdots y_{n+1}$ is linearly dependent. Since, the choice of elements were arbitrary, we see that $\mathcal{R}(T)$ has no linearly indepent set of cardinality greater than n. Hence, $\dim \mathcal{R}(T) \leq n$.

(c) Let $x_1, x_2 \in \mathcal{N}(T)$. We want to show that $\alpha x_1 + \beta x_2 \in \mathcal{N}(T)$ for any scalars α, β. Since $x_1, x_2 \in \mathcal{N}(T)$, we have $Tx_1 = 0, Tx_2 = 0$. Then,

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = 0$$

Therefore, $\alpha x_1 + \beta x_2 \in \mathcal{N}(T)$. Hence, $\mathcal{N}(T)$ is a vector space.

■

When, $T : \mathcal{D}(T) \to Y$ is injective linear operator, that is

$$Tx_1 = Tx_2 \implies x_1 = x_2 \tag{2.52}$$

then there exists an inverse map $T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ defined by $T^{-1}y = x$ where $y = Tx$.

Theorem 2.17 (Inverse Operator)
Let X, Y be vector spaces, both real or both complex. Let $T : \mathcal{D}(T) \to Y$ be a linear operator with domain $\mathcal{D}(T) \subset X$. Then:

(a) The inverse $T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ exists if and only if $Tx = 0_Y$ implies $x = 0_X$.

(b) If T^{-1} exists, then T^{-1} is a linear operator.

(c) If $\dim \mathcal{D}(T) = n < \infty$ and T^{-1} exists, then $\dim \mathcal{R}(T) = \dim \mathcal{D}T$.

Proof. (a) Suppose that $Tx = 0_Y$ implies $x = 0_X$. Let $x_1, x_2 \in \mathcal{D}(T)$ such that $Tx_1 = Tx_2$. Then,

$$T(x_1 - x_2) = Tx_1 - Tx_2 = 0_Y$$

Hence, $x_1 - x_2 = 0_X \implies x_1 = x_2$. Therefore, T is injective. Hence, by equation (2.52), T^{-1} exists.

Conversely, suppose that T^{-1} exists. Then T is injective and equation (2.52) holds. Then $Tx = 0_Y = T0_X \implies x = 0_X$.

(b) Assume that T^{-1} exists. The domain of T^{-1}, $\mathcal{R}(T)$ is a vector space by theorem (2.16)-(a). Let $y_1, y_2 \in \mathcal{R}(T)$ and α be any scalar. We want to show that $T^{-1}(\alpha y_1 + y_2) = \alpha T^{-1}y_1 + T^{-1}y_2$.

Since, $y_1, y_2 \in \mathcal{R}(T)$, we have $y_1 = Tx_1, y_2 = Tx_2$ for some $x_1, x_2 \in \mathcal{D}(T)$. Then, $x_1 = T^{-1}y_1, x_2 = T^{-1}y_2$. Hence, $\alpha T^{-1}y_1 + T^{-1}y_2 = \alpha x_1 + x_2 \in \mathcal{D}(T)$. We have,

$$\begin{aligned} T(\alpha T^{-1}y_1 + T^{-1}y_2) &= T(\alpha x_1 + x_2) \\ &= \alpha Tx_1 + Tx_2 \\ &= \alpha y_1 + y_2 \\ &= T(T^{-1}(\alpha y_1 + y_2)) \end{aligned}$$

Since, T is injective, by equation (2.52), we have $T^{-1}(\alpha y_1 + y_2) = \alpha T^{-1}y_1 + T^{-1}y_2$. Therefore, T^{-1} is a linear operator.

(c) Assume that $\dim \mathcal{D}(T) = n < \infty$. Then by theorem (2.16)-(b), $\dim \mathcal{R}(T) \leq n$. Let $\dim \mathcal{R}(T) = m$. Then, $m \leq n$. Since T^{-1} exists, and $\dim \mathcal{D}(T^{-1}) = \dim \mathcal{R}(T) = m < \infty$, we have $\dim \mathcal{R}(T^{-1}) = \dim \mathcal{D}(T) \leq m$. Hence $n \leq m$. Therefore, $m = n$ and $\dim \mathcal{R}(T) = \dim \mathcal{D}(T)$.

■

§2.7 Bounded and Continuous Linear Operator

Definition 2.7 (Bounded Linear Operator)**.** Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T : \mathcal{D}(T) \to Y$ be a linear operator where $\mathcal{D}(T) \subseteq X$. The operator is said to be bounded if there exists a real number c such that for all $x \in \mathcal{D}(T)$, we have

$$\|Tx\|_Y \leq c\,\|x\|_X \tag{2.53}$$

To emphasize that we are dealing with normed spaces, we write the map as $T : (\mathcal{D}(T), \|\cdot\|_{\mathcal{D}(T)}) \to (Y, \|\cdot\|_Y)$. When there is no source of confusion, we simply write $T : \mathcal{D}(T) \to Y$.

What is the smallest possible c that satisfies (2.53)?This is precisely $\sup_{x\in\mathcal{D}(T)\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X}$. Hence, we define:

Definition 2.8 (Operator Norm)**.** The operator norm of a bounded linear operator $T : \mathcal{D}(T) \to Y$ is defined as

$$|||T||| = \sup_{x\in\mathcal{D}(T)\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X} \tag{2.54}$$

Putting $c = |||T|||$ in equation (2.53), we get,

$$\|Tx\|_Y \leq |||T|||\,\|x\|_X \tag{2.55}$$

Lemma 2.18 (Norm)

Let $T : \mathcal{D}(T) \to Y$ be a bounded linear operator. Then:

(a)

$$\|\|T\|\| = \sup_{\|x\|_X = 1} \|Tx\|_Y \tag{2.56}$$

(b) $\|\|T\|\|$ defined in 2.8 satisfies (N1) to (N4) in definition 2.1.

Proof. (a) Let $x = \frac{y}{\|y\|_X}$. Then $\|x\|_X = \left\| \frac{y}{\|y\|_X} \right\| = \frac{1}{\|y\|_X} \|y\|_X = 1$. Hence,

$$\begin{aligned}\|\|T\|\| &= \sup_{y \in \mathcal{D}(T) \setminus \{0\}} \frac{\|Ty\|_Y}{\|y\|_X} \\ &= \sup_{y \in \mathcal{D}(T) \setminus \{0\}} \left\| \frac{1}{\|y\|_X} Ty \right\|_Y \\ &= \sup_{\|x\|_X = 1} \|Tx\|_Y\end{aligned}$$

(b)

$$\begin{aligned}\|\|T\|\| &= \sup_{x \in \mathcal{D}(T) \setminus \{0\}} \frac{\|Tx\|_Y}{\|x\|_X} \\ &\geq \sup_{x \in \mathcal{D}(T) \setminus \{0\}} 0 \\ &\geq 0\end{aligned}$$

Hence, (N1) is satisfied.

Let $Tx = 0$ for all x. Then $\|Tx\|_Y = 0$ for all x. Hence, $\|\|T\|\| = 0$. Conversely, if $\|\|T\|\| = 0$, then $\sup_{\|x\|=1} \|Tx\|_Y = 0$. Hence, $\|Tx\|_Y = 0$ for all x. Hence, $Tx = 0$ for all x. Hence, (N2) is satisfied.

Let α be any scalar. Then (N3) is satisfied because:

$$\begin{aligned}\|T(\alpha x)\|_Y &= \sup \|x\|_X = 1 \, \|T(\alpha x)\|_Y \\ &= \sup \|x\|_X = 1 |\alpha| \, \|Tx\|_Y \\ &= |\alpha| \sup \|x\|_X = 1 \, \|Tx\|_Y \\ &= |\alpha| \|\|T\|\|\end{aligned}$$

Let T_1, T_2 both be bounded linear operators from $\mathcal{D}(T) \to Y$. Then,

$$\begin{aligned}\|\|T_1 + T_2\|\| &= \sup_{\|x\|=1} \|(T_1 + T_2)x\|_Y \\ &= \sup_{\|x\|=1} \|T_1 x + T_2 x\|_Y \\ &\leq \sup_{\|x\|=1} \|T_1 x\|_Y + \|T_2 x\|_Y \\ &= \sup_{\|x\|=1} \|T_1 x\|_Y + \sup_{\|x\|=1} \|T_2 x\|_Y \\ &= \|\|T_1\|\| + \|\|T_2\|\|\end{aligned}$$

Hence, (N4) is satisfied. Therefore, $\|\|T\|\|$ defined in 2.8 satisfies (N1) to (N4) in Definition 2.1. ■

§2.7.i Examples of Bounded Linear Operators

Example 2.6 (Identity Operator)**.** Let $(X, \|\cdot\|_X)$ be a normed space. Then the identity operator $I : X \to X$ defined by $Ix = x$ is a bounded linear operator. If $X \neq \{0\}$, then $\|Ix\|_X = \sup_{\|x\|=1} \|x\|_X$. Hence, $|||I||| = 1$.

Example 2.7 (Zero Operator)**.** The zero operator $\hat{0} : X \to Y$ defined by $\hat{0}x = 0, \forall x \in X$ is bounded and $|||0||| = 0$.

Example 2.8 (Differentiation Operator)**.** Let $(\rho[0,1], \|\cdot\|_\rho$ be the normed space of all complex valued polynomials on $J = [0,1]$ and $\|\cdot\|_\rho$ be the max norm given by,

$$|||x|||_\rho = \max_{t\in J} |x(t)| \tag{2.57}$$

A differentatioin operator $D : (\rho[0,1], \|\cdot\|_\rho) \to (\rho[0,1], \|\cdot\|_\rho)$ is defined by $Dx(t) = x'(t)$. Then D is linear due to the linearity of differentiation. We claim that D is not bounded. Let $x_n(t) = t^n$ with $n \in \mathbb{N}$. Then,

$$|||x_n|||_\rho = \max_{t\in J} |t^n| = 1$$

and $Dx_n(t) = nt^{n-1}$. Hence,

$$\|Dx_n\|_\rho = \max_{t\in J} |nt^{n-1}| = n$$

Then we have,

$$\begin{aligned} |||D||| &= \sup_{\|x\|_\rho=1} \|Dx\|_\rho \\ &\geq \frac{\|Dx_n\|_\rho}{\|x_n\|_\rho} \\ &= \frac{n}{1} = n \end{aligned}$$

Hence, D is not bounded.

Example 2.9 (Matrix)**.** A real matrix $A = (\alpha_{jk})_{j=1,\ldots,r;k=1,\ldots,n}$ with r rows and n columns defines an operator $A : \mathbb{R}^n \to \mathbb{R}^r$ by means of

$$y = Ax \tag{2.58}$$

where $x = (\xi_j)_{j=1,\ldots,n}$ and $y = (\eta_j)_{j=1,\ldots,r}$ are column vectors with n and r components respectively. In terms of components, 2.58 becomes,

$$\eta_j = \sum_{k=1}^{n} \alpha_{jk}\xi_k, (j = 1, \ldots, r) \tag{2.59}$$

The operator A is linear because matrix multiplication is linear operation. Let us prove that A is bounded. Recall that the norm on $\mathbb{R}^n$ is given by,

$$\|x\|_{\mathbb{R}^n} = \left(\sum_{m=1}^{n} |\xi_m|^2\right)^{1/2} \tag{2.60}$$

And the norm of y reads,

$$\|y\|_{\mathbb{R}^r} = \left(\sum_{j=1}^{r} |\eta_j|^2\right)^{1/2} \tag{2.61}$$

By applying the Cauchy-Schwarz inequality to 2.59, one obtains,

$$
\begin{aligned}
\|Ax\|_{\mathbb{R}^r}^2 &= \|y\|_{\mathbb{R}^r}^2 \\
&= \sum_{j=1}^{r} |\eta_j|^2 \\
&= \sum_{j=1}^{r} \left(\sum_{k=1}^{n} \alpha_{jk}\xi_k \right)^2 \\
&\leq \sum_{j=1}^{r} \left(\sum_{k=1}^{n} |\alpha_{jk}\xi_k| \right)^2 \\
&\leq \sum_{j=1}^{r} \left[\left(\sum_{k=1}^{n} |\alpha_{jk}|^2 \right) \left(\sum_{m=1}^{n} |\xi_m|^2 \right) \right] \\
&= \left(\sum_{m=1}^{n} |\xi_m|^2 \right) \left(\sum_{j=1}^{r} \sum_{k=1}^{n} \alpha_{jk}^2 \right) \\
&= \|x\|_{\mathbb{R}^n}^2 \sum_{j=1}^{r} \sum_{k=1}^{n} \alpha_{jk}^2 \\
&= \|x\|_{\mathbb{R}^n}^2 \, c^2 \quad \text{where} \quad c = \left(\sum_{j=1}^{r} \sum_{k=1}^{n} \alpha_{jk}^2 \right)^{1/2}
\end{aligned}
$$

Hence, $\|Ax\|_{\mathbb{R}^r} \leq c \, \|x\|_{\mathbb{R}^n}$. Therefore, A is bounded linear operator.

Theorem 2.19 (Finite Dimension)
If a normed space X is finite dimensional, then every linear opeartor on X is bounded.

Proof. Let $T : X \to Y$ be any linear operator. Let $dim X = n$ and $\{e_1, \dots, e_n\}$ be a basis for X. Let $x \in X$, then $x = \sum_{j=1}^{n} \xi_j e_j$. Then,

$$
\begin{aligned}
\|Tx\| &= \left\| \sum_{j=1}^{n} \xi_j T e_j \right\| \\
&\leq \sum_{j=1}^{n} |\xi_j| \, \|T e_j\| \\
&\leq \max_{k=1}^{n} \|T e_k\| \sum_{j=1}^{n} |\xi_j|
\end{aligned}
$$

Then using Lemma 2.2 with $\alpha_j = \xi_j$ and $x_j = e_j$,

$$
\sum_{j=1}^{n} |\xi_j| = \frac{1}{c} \left\| \sum_{j=1}^{n} \xi_j \right\| = \frac{1}{c} \|x\|
$$

Let $d = \frac{1}{c} \max_{k=1}^{n} \|T e_k\|$. Then we have,

$$
\|Tx\| \leq d \, \|x\|
$$

Hence, T is bounded. ■

Theorem 2.20 (Continuity and Boundedness)
Let $T : \mathcal{D}(T) \to Y$ be a linear operator where $\mathcal{D}(T) \subseteq X$ and X, Y are normed spaces. Then:

(a) T is continuous if and only if T is bounded.

(b) If T is continuous at a single point, it is continuous everywhere.

Proof. (a) For $T = \hat{0}$, T is both bounded and continuous. Let $T \neq \hat{0}$ be bounded. Let $x_0 \in \mathcal{D}(T)$ and $\epsilon > 0$ be arbitrary. Let $\delta = \frac{\epsilon}{\|T\|}$. Then for all $x \in \mathcal{D}(T)$ such that $\|x - x_0\| < \delta$, we have,

$$\begin{aligned}\|Tx - Tx_0\| &= \|T(x - x_0)\| \\ &\leq \|T\| \, \|x - x_0\| \\ &< \|T\|\delta \\ &= \epsilon\end{aligned}$$

Therefore, T is continuous at x_0. Since, x_0 was arbitrary, T is continuous everywhere.

Conversely, assume that T is continuous. Choose an arbitrary $x_0 \in \mathcal{D}(T)$. Then for any $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in \mathcal{D}(T)$ such that $\|x - x_0\| < \delta$, we have $\|Tx - Tx_0\| < \epsilon$.

Let $y \in \mathcal{D}(T) \setminus \{0\}$ and set $x = x_0 + \frac{\delta}{\|y\|}y$. Then, $x - x_0 = \frac{\delta}{\|y\|}y$. Hence, $\|x - x_0\| = \frac{\delta}{\|y\|}\|y\| = \delta$. Therefore, $\|Tx - Tx_0\| < \epsilon$. Hence,

$$\begin{aligned}&\|Tx - Tx_0\| < \epsilon \\ \implies &\|T(x - x_0)\| < \epsilon \\ \implies &\left\|T\left(\frac{\delta}{\|y\|}y\right)\right\| < \epsilon \\ \implies &\frac{\delta}{\|y\|}\|Ty\| < \epsilon \\ \implies &\|Ty\| < \frac{\epsilon}{\delta}\|y\| \\ \implies &\|Ty\| < c\,\|y\|\end{aligned}$$

Where $c = \frac{\epsilon}{\delta}$. Therefore, T is bounded.

(b) Let T be continuous at a point $x_0 \in \mathcal{D}(T)$. Then using the second argument of part (a), T is bounded. Then using part(a) again, T is continuous everywhere.

■

Corollary 2.21 (Continuity, null space)
Let T be a bounded linear operator. Then:

(a) $x_n \to x$ where $x_n \in \mathcal{D}(T)$ implies $Tx_n \to Tx$.

(b) The null space $\mathcal{N}(T)$ is closed.

Proof. (a) Since T is bounded, T is continuous by Theorem 2.20-(a). Hence, $x_n \to x$ implies $Tx_n \to Tx$.

(b) Let (x_n) be a sequence in $\mathcal{N}(T)$ such that $x_n \to x$. Then $Tx_n = 0$ for all n. Hence $Tx_n \to 0$. Since T is continuous, $Tx_n \to Tx$. Hence, $Tx = 0$. Therefore, $x \in \mathcal{N}(T)$. Hence, $\mathcal{N}(T)$ is closed. ■

Lemma 2.22

Let $T_2 : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$, $T_1 : (Y, \|\cdot\|_Y) \to (Z, \|\cdot\|_Z)$ and $T : (X, \|\cdot\|_X) \to (X, \|\cdot\|_X)$ be bounded linear operators. Show that

(a) $|||T_1 T_2||| \le |||T_1||| \cdot |||T_2|||$

(b) $|||T^n||| \le |||T|||^n$ for any positive integer n

Proof. (a) Let $x \in X$, then

$$\begin{aligned}
|||T_1 T_2||| &= \sup_{x_0 \in X} \frac{\|T_1 T_2 x_0\|_Z}{\|x_0\|_X} \\
&\le \frac{\|T_1 T_2 x\|_Z}{\|x\|_X} \\
&\le \frac{|||T_1||| \cdot \|T_2 x\|_Y}{\|x\|_X} \\
&\le \frac{|||T_1||| \cdot |||T_2||| \cdot \|x\|_X}{\|x\|_X} \\
&= |||T_1||| \cdot |||T_2|||
\end{aligned}$$

(b) For $n = 1$, the claim is trivially true. Let us assume that the claim is true for some $n = k \in \mathbb{N}$ where $k \ge 1$. Then,

$$\begin{aligned}
\left|\left|\left|T^{k+1}\right|\right|\right| &= \left|\left|\left|T^k T\right|\right|\right| \\
&\le \left|\left|\left|T^k\right|\right|\right| \cdot |||T||| \quad \text{by part (a)} \\
&\le |||T|||^k \cdot |||T||| \\
&= |||T|||^{k+1}
\end{aligned}$$

Therefore, by induction, the claim is true for all positive $n \in \mathbb{N}$. ■

Definition 2.9 (Equality, Restriction and Extension). Two operators T_1 and T_2 are said to be **equal**, written as $T_1 = T_2$ if they have the same domain, $\mathcal{D}(T_1) = \mathcal{D}(T_2)$ and for all $x \in \mathcal{D}(T_1)$, we have $T_1 x = T_2 x$.

The **restriction** of an operator $T : \mathcal{D}(T) \to Y$ to a subset $B \subseteq \mathcal{D}(T)$ is denoted by $T|_B$ and is the operator defined by,

$$T|_B : B \to Y, \quad T|_B x = Tx, \quad \forall x \in B$$

An **extension** of T to a set $M \supseteq \mathcal{D}(T)$ is an operator,

$$\hat{T} : M \to Y, \quad \text{such that} \quad \hat{T}|_{\mathcal{D}(T)} = T$$

Theorem 2.23 (Bounded Linear Extension)

Let $T : \mathcal{D}(T) \to Y$ be a bounded linear operator where $\mathcal{D}(T) \subseteq X$ and $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$

are normed spaces. Then T has an extension

$$\hat{T} : \overline{\mathcal{D}(T)} \to Y$$

which is bounded and $\left\|\left|\hat{T}\right|\right\| = |||T|||$

Proof. Let $x \in \overline{\mathcal{D}(T)}$. Then there exists a sequence (x_n) in $\mathcal{D}(T)$ such that $x_n \to x$. Since T is bounded and linear,

$$\begin{aligned} \|Tx_n - Tx_m\| &= \|T(x_n - x_m)\| \\ &\leq |||T|||\, \|x_n - x_m\| \end{aligned}$$

Since, $\|x_n - x_m\| \to 0$ as $n, m \to \infty$, the sequence (Tx_n) is a Cauchy sequence in Y. Since Y is banach, it is complete, hence (Tx_n) converges to some $y \in Y$. Define $\hat{T} : \overline{\mathcal{D}(T)} \to Y$ by $\hat{T}x = y$. We need to check well definedness of $\hat{T}$.

Let (z_n) be another sequence in $\mathcal{D}(T)$ such that $z_n \to x$. Consider (v_n) be the sequence $(x_1, z_1, x_2, z_2, \ldots)$. Then (v_n) converges to x. Hence, the sequence Tv_n converges to some $y' \in Y$. Since, every subsequences of a convergent sequence converges to the same limit, and Tx_n is a subsequence of Tv_n, we have $y = y'$. Since, Tz_n is also a subsequence of Tv_n, Tz_n also converges to y. Hence, $\hat{T}$ is well defined.

Let $x_1, x_2 \in \overline{\mathcal{D}(T)}$ and α be any scalar. Then,

$$\begin{aligned} \hat{T}(\alpha x_1 + x_2) &= \lim_{n\to\infty} T(\alpha x_1 + x_2) \\ &= \lim_{n\to\infty} \alpha T x_1 + T x_2 \\ &= \alpha \lim_{n\to\infty} T x_1 + \lim_{n\to\infty} T x_2 \\ &= \alpha \hat{T} x_1 + \hat{T} x_2 \end{aligned}$$

Therefore, $\hat{T}$ is linear. For every $x \in \mathcal{D}(T)$, we have $\hat{T}x = Tx$. Hence, $\hat{T}$ is an extension of T.

$\|Tx_n\| \leq |||T|||\, \|x_n\|$ for all n. Since, $\|.\|$ is a continuous mapping, we have after taking limit on both sides as $n \to \infty$,

$$\begin{aligned} &\left\|\hat{T}x\right\| \leq |||T|||\, \|x\| \\ \implies &\sup_{\|x\|=1} \left\|\hat{T}x\right\| \leq |||T||| \\ \implies &\left\|\left|\hat{T}\right|\right\| \leq |||T||| \end{aligned}$$

Hence, $\hat{T}$ is bounded. Moreover, $\left\|\left|\hat{T}\right|\right\| \geq |||T|||$ because $\hat{T}$ is an extension of T. Therefore, $\left\|\left|\hat{T}\right|\right\| = |||T|||$. ■

§2.8 Linear Functional

Definition 2.10 (Linear Functional)**.** A linear functional f is a linear operator with domain in a vector space X and codomain in the field $\mathbb{K}$ of X, that is

$$f : X \to \mathbb{K}$$

where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Definition 2.11 (Bounded Linear Functional). A bounded linear functional f is a bounded linear operator with domain in a normed space $(X, \|\cdot\|)$ and codomain in the field $\mathbb{K}$ of X, such that there exists a real number c such that for all $x \in X$,

$$|f(x)| \leq c \|x\|$$

Furthermore, the norm of f is defined as in 2.8,

$$|||f||| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)| \tag{2.62}$$

And using 2.55, we can write

$$|f(x)| \leq |||f||| \, \|x\| \tag{2.63}$$

Theorem 2.24 (Countinuity and Boundedness)
A linear functional f is continuous if and only if it is bounded.

Proof. This is a special case of 2.20. ■

Example 2.10. Consider the normed space $(C[a,b], \|\cdot\|_{\max})$ with respect to the max norm

$$\|x\|_{\max} = \max_{t \in [a,b]} |x(t)|$$

with $x : [a,b] \to \mathbb{C}$ being a complex valued continuous function on $[a,b]$. Fix $t_0 \in J = [a,b]$ and define a linear functional f on all of $(C[a,b], \|\cdot\|_{\max})$, i.e $f : (C[a,b], \|\cdot\|_{\max}) \to \mathbb{C}$ as,

$$f(x) = x(t_0)$$

where $x \in C[a,b]$. We will show that f is a bounded linear functional. Let $x, y \in C[a,b]$ and $\alpha \in \mathbb{C}$, then

$$f(\alpha x + y) = (\alpha x + y)(t_0) = \alpha x(t_0) + y(t_0) = \alpha f(x) + f(y)$$

Hence, f is linear. $\|f(x)\| = \|x(t_0)\| \leq \|x\|_{\max}$. Therefore, f is bounded. Moreover,

$$\begin{aligned} &\frac{\|f(x)\|}{\|x\|_{\max}} \leq 1 \\ \implies &\sup_{\|x\|_{\max}=1} \|f(x)\| \leq 1 \\ \implies &|||f||| \leq 1 \end{aligned}$$

Let $x_0 \in C[a,b]$ be the constant function such that $x_0(t) = 1$ for all $t \in [a,b]$. Then,

$$|||f||| \geq \|f(x_0)\| = \|x_0(t_0)\| = 1$$

Therefore, $|||f||| = 1$.

Example 2.11. Consider the hilbert space l^2 (see 1.7). We define a linear functional f on l^2 by fixing $a = (\alpha_n) \in l^2$ and defining f as,

$$f(x) = \sum_{n=1}^{\infty} \alpha_n \xi_n \tag{2.64}$$

where $x = (\xi_n) \in l^2$. The series in (2.64) converges absolutely. Then using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\|f(x)\| &= \left|\sum_{n=1}^{\infty} \alpha_n \xi_n\right| \\
&\leq \sum_{n=1}^{\infty} |\alpha_n \xi_n| \\
&\leq \left(\sum_{n=1}^{\infty} |\alpha_n|^2\right)^{1/2} \left(\sum_{n=1}^{\infty} |\xi_n|^2\right)^{1/2} \\
&= \|a\| \, \|x\|
\end{aligned}$$

Therefore, f is a bounded linear functional on l^2.

Definition 2.12 (Algebraic Dual Space)**.** The algebraic dual space of a vector space X is the set of all linear functionals on X and is denoted by X^*. That is,

$$X^* = \{f : X \to \mathbb{K} \mid f \text{ is linear}\}$$

Let $f, g \in X^*$ and $\alpha \in \mathbb{K}$, then $f + g$ and αf are defined as,

$$\begin{aligned}
(f + g)(x) &= f(x) + g(x) \\
(\alpha f)(x) &= \alpha f(x)
\end{aligned}$$

The algebraic dual space $(X^*)^*$ of X^*, denoted as X^{**}, is called the second algebraic dual space of X. We can define an element of X^{**}, g for a fixed $x \in X$ as,

$$\begin{aligned}
&g : X^* \to \mathbb{K} \\
&g(f) = g_x(f) = f(x) \quad \forall f \in X^*
\end{aligned}$$

The linearity of g can be seen from,

$$g_x(\alpha f + g) = (\alpha f + g)(x) = \alpha f(x) + g(x) = \alpha g_x(f) + g_x(g)$$

for all $f, g \in X^*$ and $\alpha \in \mathbb{K}$. Hence, each $x \in X$ corresponds to a $g_x \in X^{**}$. This defined the following mapping:

Definition 2.13 (Cannonical mapping or Cannonical Embedding)**.** The cannonical mapping or cannonical embedding of a vector space X into its second algebraic dual space X^{**} is defined as,

$$C : X \to X^{**} \qquad\qquad C(x) = g_x \quad \forall x \in X$$

Let $x, y \in X$ and $\alpha \in \mathbb{K}$, then for all $f \in X^*$,

$$C(\alpha x + y)(f) = f(\alpha x + y) = \alpha f(x) + f(y) = \alpha C(x)(f) + C(y)(f)$$

Therefore, C is linear. Assume that $C(x) = C(y)$ then for all $f \in X^*$,

$$\begin{aligned}
&C(x) = C(y) \\
&\implies g_x = g_y \\
&\implies g_x(f) = g_y(f) \\
&\implies f(x) = f(y)
\end{aligned}$$

Since, this is true for all f, we have $x = y$. Therefore, C is injective.

Definition 2.14 (Isomorphism of Normed Space)**.** Two normed spaces X and Y are said to be isomorphic if there exists a bijective linear operator $T : X \to Y$ such that both T and T^{-1} are isometric. That is,

$$\begin{aligned}
&T : X \to Y \quad \text{and} \quad T^{-1} : Y \to X \\
&T(x+y) = Tx + Ty \quad \forall x, y \in X \\
&T(\alpha x) = \alpha T(x) \quad \forall x \in X, \alpha \in \mathbb{K} \\
&\|Tx\|_Y = \|x\|_Y \quad \forall x \in X \\
&\left\|T^{-1}y\right\|_X = \|y\|_X \quad \forall y \in Y
\end{aligned}$$

Definition 2.15 (Embeddable)**.** A vector space X is said to be embeddable in a vector space Y if X is (vector space) isomorphic to a subspace of Y.

Since, the cannonical map C is injective, X is (vector space) isomorphic to the range of C in X^{**}. Therefore, X is embeddable in X^{**}. If X is such a space so that C is also surjective, that is Cc is bijective linear operator, then X is isomorphic to X^{**}.

Definition 2.16 (Algebraic Reflexivity)**.** A vector space X is said to be algebraically reflexive if it is (vector space) isomorphic to its second algebraic dual space X^{**}.

Theorem 2.25 (Dimension of X^*)

Let X be an n-dimensional vector space and $E = \{e_1, e_2, \dots, e_n\}$ be a basis of X. Then, $F = \{f_1, f_2, \dots, f_n\}$ defined as,

$$f_i(e_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

is a basis of X^* and $\dim X^* = \dim X = n$.

Proof. Let β_i be a set of scalars such that,

$$\sum_{i=1}^{n} \beta_i f_i(x) = 0$$

Then, for a fixed $j \in \{1, 2, \dots, n\}$,

$$\sum_{i=1}^{n} \beta_i f_i(e_j) = \sum_{i=1}^{n} \beta_i \delta_{ij} = \beta_j = 0$$

Therefore, $\beta_i = 0$ for all $i \in \{1, 2, \dots, n\}$. Hence, F is linearly independent.

Let $f \in X^*$. Since f is linear, we can write $f(e_i) = \alpha_i$ for some $\alpha_i \in \mathbb{K}$. Then,

$$f(x) = \sum_{i=1}^{n} \xi_i \alpha_i$$

Also,

$$f_i(x) = f_i\left(\sum_{i=1}^{n} \xi_i e_i\right) = \sum_{i=1}^{n} \xi_i f_i(e_i) = \xi_i$$

Hence, we have $f(x) = \sum_{i=1}^{n} \alpha_i f_i(x)$. Therefore, F spans X^*, hence F is a basis of X^*. Therefore, $\dim X^* = n = \dim X$. ∎

Lemma 2.26 (Zero Vector)
Let X be a finite dimensional vector space. Let $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in X^*$. Then, $x_0 = 0$.

Proof. Let $\{e_1, e_2, \dots, e_n\}$ be a basis of X. Then $x_0 = \sum_{i=1}^n \xi_i e_i$. Let $f \in X^*$, then

$$f(x_0) = \sum_{i=1}^{n} \xi_i \alpha_i = 0$$

This is 0 by assumption for all f, hence also for every choice of α_i. Therefore, ξ_i must be 0 for all i. Hence, $x_0 = 0$. ∎

Theorem 2.27 (Algebraic Reflexivity of Finite Dimensional Space)
A finite dimensional vector space X is algebraically reflexive.

Proof. Consider the cannonical mapping $C : X \to X^{**}$ defined in 2.13. Assume that $C(x_0) = 0$ for some $x_0 \in X$. Then for all $f \in X^*$,

$$(Cx_0)(f) = g_x(f) = f(x_0) = 0$$

Then by lemma 2.26, $x_0 = 0$. Hence, by theorem 2.17, C has an inverse $C^{-1} : \mathcal{R}(C) \to X$ and $\dim \mathcal{R}(C) = \dim X$. By theorem 2.25, $\dim X^{**} = \dim X = \dim \mathcal{R}(C)$. Hence, $\mathcal{R}(C)$ is not a proper subspace of X^{**}. Therefore C is surjective and X is algebraically reflexive. ∎

§2.9 Normed spaces of operators

Take any two normed spaces X and Y (both over real or complex numbers) and consider the set

$$\mathcal{B}(X, Y)$$

consisting of all bounded linear operators from X to Y. Each such operator is defined over all of X and its range is contained in Y. We first want to see that $\mathcal{B}(X, Y)$ itself can be made into a normed space.

Definition 2.17 (Vector Space Structure)**.** Vector addition is defined as

$$(T_1 + T_2)x = T_1 x + T_2 x \tag{2.65}$$

where $T_1, T_2 \in \mathcal{B}(X, Y)$.
For any scalar α, the scalar multiplication is defined by

$$(\alpha T)x = \alpha(Tx) \tag{2.66}$$

$\forall T \in \mathcal{B}(X, Y)$.

Definition 2.18 (Bounded Linear Operators form a Normed Space)**.** The Vector Space $\mathcal{B}(X, Y)$ of all bounded linear operators from a normed space $(X, \|\cdot\|_X)$ to a normed space $(Y, \|\cdot\|_Y)$ is itself a normed space with norm defined by

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|=1} \|Tx\|_Y \tag{2.67}$$

In what case $\mathcal{B}(X, Y)$ will be a Banach space? The question is answered in the following theorem:

Theorem 2.28 (The Space of Bounded Linear Operator is Banach)
If $(Y, \|\cdot\|_Y)$ is a Banach space, then $(\mathcal{B}(X, Y), |||\cdot|||)$ is a Banach space.

Proof. Consider an arbitrary Cauchy sequence (T_n) in $\mathcal{B}(X, Y)$ and show that $T_n \to T$ in $\mathcal{B}(X, Y)$. Since (T_n) is Cauchy, $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that

$$\|T_n - T_m\| < \epsilon, \quad \text{whenever } m, n > N \tag{2.68}$$

Then for all $x \in X$ and $m, n > N$, one has

$$\begin{aligned} \|T_n x - T_m x\|_Y &= \|(T_n - T_m)x\|_Y \\ &\leq \|T_n - T_m\| \|x\|_X \\ &< \epsilon \|x\|_X \quad [\text{By } (2.68)] \end{aligned} \tag{2.69}$$

$T_n x - T_m x \in Y \ \forall n \in \mathbb{N}$.

Now, fix $x \in X$. Given $\tilde{\epsilon} > 0$, we may choose $\epsilon = \frac{\tilde{\epsilon}}{\|x\|_X}$ in (2.68) so that

$$\epsilon \|x\|_X < \tilde{\epsilon} \tag{2.70}$$

Then from (2.69), one obtains, $\|T_n x - T_m x\|_Y < \tilde{\epsilon}$ and from (2.68), one finds that $(T_n x)$ is a Cauchy sequence in $(Y, \|\cdot\|_Y)$. Since $(Y, \|\cdot\|_Y)$ is complete, $(T_n x)$ converges in Y, say, $T_n x \to y$. We thus find a vector $y \in Y$ from a vector $x \in X$ with the help of the Cauchy sequence (T_n) in $\mathcal{B}(X, Y)$. In other words, one has an operator $T : X \to Y$ defined by

$$y = \lim_{n\to\infty} T_n x =: Tx \tag{2.71}$$

The operator $T : X \to Y$ is linear since

$$\begin{aligned} T(\alpha x + \beta z) &= \lim_{n\to\infty} T_n(\alpha x + \beta z) = \alpha \lim_{n\to\infty} T_n x + \beta \lim_{n\to\infty} T_n z \\ &= \alpha Tx + \beta Tz \end{aligned}$$

$\forall x, z \in X$ and α, β scalars.

Now, let us prove that $T : X \to Y$ defined in (2.71) is bounded. In (2.68), we saw that $\forall \epsilon > 0 \exists n \in \mathbb{N}$ such that $\forall x \in X$ and $m, n > N$,

$$\|T_n x - T_m x\|_Y < \epsilon \|x\|_X \tag{2.72}$$

From the above, we see that for every $n > N$ and all $x \in X$

$$\begin{aligned} \|T_n x - Tx\|_Y &= \|T_n x - \lim_{m\to\infty} T_m x\|_Y \\ &= \lim_{m\to\infty} \|T_n x - T_m x\|_Y \quad [\|\cdot\|_Y \text{ is a continuous function on } Y] \\ &< \epsilon \|x\|_X \quad [\text{By } (2.72)] \end{aligned}$$

One, therefore, finds that for all $x \in X$ with $n > N$

$$\|(T_n - T)x\|_Y \leq \epsilon \|x\|_X \tag{2.73}$$

which shows that $T_n - T$ is a bounded linear operator for $n > N$. Since T_n is bounded for $\forall n \in \mathbb{N}$ $[T_n \in \mathcal{B}(X, Y)]$, one must have $T_n - (T_n - T) = T$ is also bounded. In other words, T is a bounded linear operator, i.e., $T \in \mathcal{B}(X, Y)$. Take supremum over all $x \in X$ with $\|x\|_X = 1$ on both sides of (2.73) to obtain the following

$$\begin{aligned} \sup_{\substack{x\in X \\ \|x\|_X = 1}} \|(T_n - T)x\|_Y &\leq \epsilon \\ \implies \|T_n - T\| &\leq \epsilon \end{aligned} \tag{2.74}$$

(2.74) means that $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, s.t. $\forall n > N$, one has $\|T_n - T\| < \epsilon$. Hence, $\|T_n - T\| \to 0$. In other words, $T_n \to T$ in $\mathcal{B}(X, Y)$. ∎

Definition 2.19 (Dual Space)**.** Let $(X, \|\cdot\|_X)$ be a normed space. Then the set of all bounded linear functionals on X constitutes a normed space with norm defined by

$$\|f\| = \sup_{\substack{x\in X\\ x\neq 0}} \frac{|f(x)|}{\|x\|_X} = \sup_{\|x\|=1} |f(x)| \tag{2.75}$$

which is called the dual space of X and is denoted by X'.

Since a linear functional on a normed space $(X, \|\cdot\|_X)$ is a map from $(X, \|\cdot\|_X)$ to $\mathbb{R}$ or $\mathbb{C}$ (depending on if X is a real or a complex vector space), we see that X' is simply $\mathcal{B}(X, \mathbb{K})$ with the complete space $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ (with the Euclidean or unitary metric). Hence, Theorem 2.28 applies and we have the following:

Theorem 2.29 (Dual of a Normed Space is a Banach Space)

The dual space X' of a normed space $(X, \|\cdot\|_X)$ is a Banach space (whether or not X is).

Definition 2.20. An **isomorphism** of a normed space $(X, \|\cdot\|_X)$ to a normed space $(\tilde{X}, \|\cdot\|_{\tilde{X}})$ is a bijective linear operator $T : (X, \|\cdot\|_X) \to (\tilde{X}, \|\cdot\|_{\tilde{X}})$ which preserves the norm, i.e., $\forall x \in X$, one has

$$\|Tx\|_{\tilde{X}} = \|x\|_X \tag{2.76}$$

If d_X is the metric on X induced by $\|\cdot\|_X$ and $d_{\tilde{X}}$ is the metric on $\tilde{X}$ induced by $\|\cdot\|_{\tilde{X}}$, then it is easy to verify that the following holds

$$d_{\tilde{X}}(Tx, Ty) = d_X(x, y) \tag{2.77}$$

$\forall x, y \in X$. In other words, $T : (X, d_X) \to (\tilde{X}, d_{\tilde{X}})$ is actually an isometry in the respective induced metrics. Also, (2.76) tells us that the bijective linear operator $T : (X, \|\cdot\|_X) \to (\tilde{X}, \|\cdot\|_{\tilde{X}})$ is actually bounded and hence continuous by Theorem 4.2.

Example 2.12 (Dual of $\mathbb{R}^n$)**.** The dual space of $(\mathbb{R}^n, \|\cdot\|_2)$ is $(\mathbb{R}^n, \|\cdot\|_2)$

Proof. By Theorem 4.1, if a normed space $(X, \|\cdot\|_X)$ is finite-dimensional, then every linear operator on $(X, \|\cdot\|_X)$ is bounded. Hence on $\mathbb{R}^n$, the dual $\mathbb{R}^{n*}$ coincides with the dual $\mathbb{R}^n$, i.e., $\mathbb{R}^{n*} = \mathbb{R}^n$. Given a basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ and f, for $f \in \mathbb{R}^n$, one has

$$\begin{aligned} x &= \sum_{k=1}^{n} \xi_k e_k \\ f(x) &= f\left(\sum_{k=1}^{n} \xi_k e_k\right) = \sum_{k=1}^{n} \xi_k f(e_k) \quad \text{[By linearity of } f \in \mathbb{R}^n\text{]} \\ &= \sum_{k=1}^{n} \xi_k \gamma_k, \quad \gamma_k = f(e_k) \end{aligned} \tag{2.78}$$

By Cauchy-Schwarz,

$$\begin{aligned} |f(x)| &= \left|\sum_{k=1}^{n} \xi_k \gamma_k\right| \leq \left(\sum_{k=1}^{n} |\xi_k|^2\right)^{1/2} \left(\sum_{m=1}^{n} |\gamma_m|^2\right)^{1/2} \\ &= \|x\|_2 \left(\sum_{m=1}^{n} |\gamma_m|^2\right)^{1/2} \end{aligned} \tag{2.79}$$

Taking supremum over all $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$ on both sides of (2.79), one obtains,

$$\begin{aligned} \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|_2 = 1}} |f(x)| &\leq \left(\sum_{m=1}^{n} |\gamma_m|^2 \right)^{1/2} \\ \implies \|f\| &\leq \left(\sum_{m=1}^{n} |\gamma_m|^2 \right)^{1/2} \end{aligned} \tag{2.80}$$

When $\xi_k = \gamma_k$, $\forall k \in \{1, \ldots, n\}$, (2.79) reads as

$$|f(x)| = \sum_{k=1}^{n} |\gamma_k|^2 = \sum_{k=1}^{n} \xi_k^2 = \|x\|_2^2 \quad \text{when } x = (\gamma_1, \ldots, \gamma_n)$$

In other words,

$$\frac{|f(x)|}{\|x\|_2} = \left(\sum_{k=1}^{n} \gamma_k^2 \right)^{1/2} \qquad \text{when } x = (\gamma_1, \ldots, \gamma_n) \tag{2.81}$$

And, from (2.79),

$$\frac{|f(x)|}{\|x\|_2} \leq \left(\sum_{k=1}^{n} \gamma_k^2 \right)^{1/2}, \quad \forall x \in \mathbb{R}^n \text{ and } x \neq 0 \tag{2.82}$$

(2.81) together with (2.82) imply

$$\begin{aligned} \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{|f(x)|}{\|x\|_2} &= \left(\sum_{k=1}^{n} \gamma_k^2 \right)^{1/2} \\ |||f||| &= \left(\sum_{k=1}^{n} \gamma_k^2 \right)^{1/2} \end{aligned} \tag{2.83}$$

(2.83) shows that the norm of f is equal to the Euclidean norm $\|c\|_2$, where $c = (\gamma_k) \in \mathbb{R}^n$, i.e.,

$$|||f||| = \|c\|_2 \tag{2.84}$$

with $c = (\gamma_k)_{k=1}^n \in \mathbb{R}^n$. Hence, the mapping $g : \mathbb{R}^{n'} \to \mathbb{R}^n$ defined by $g(f) = c = (\gamma_k)$, with $\gamma_k = f(e_k)$, $\{e_1, \ldots, e_n\}$ being a basis of $\mathbb{R}^n$ is norm preserving. Indeed, (2.84) tells us that

$$\|g(f)\|_2 = |||f|||$$

In fact, the isomorphism is $g : (\mathbb{R}^{n'}, \|\cdot\|) \to (\mathbb{R}^n, \|\cdot\|_2)$. ■

Example 2.13 (Other examples of Dual Spaces). 1. The dual space of ℓ^1 is ℓ^∞.

2. The dual space of ℓ^p is ℓ^q; here, $1 < p < \infty$ and q is the conjugate of p, i.e., $\frac{1}{p} + \frac{1}{q} = 1$.

3 Inner Product Space and Hilbert Space

§3.1 Inner Product Space

Definition 3.1 (Inner product spaces and Hilbert spaces)**.** An inner product space $(X, \langle\cdot,\cdot\rangle)$ is a vector space X with an inner product $\langle\cdot,\cdot\rangle$ defined on X. A Hilbert space is a complete inner product space (complete in the metric induced by the inner product as mentioned below). Suppose X is a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Then $\langle\cdot,\cdot\rangle$ is a mapping $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$ such that $\forall x, y, z \in X$ and $\alpha \in \mathbb{K}$, one has

(IP1) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

(IP2) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$

(IP3) $\langle x, y\rangle = \overline{\langle y, x\rangle}$

(IP4) $\langle x, x\rangle \geq 0$
$\langle x, x\rangle = 0 \Leftrightarrow x = 0_X$

An inner product $\langle\cdot,\cdot\rangle$ on X defines a norm on X given by

$$\|x\|_{in} = \sqrt{\langle x, x\rangle} \tag{3.1}$$

and a metric on X by

$$d_{in}(x, y) = \|x - y\|_{in} = \sqrt{\langle x - y, x - y\rangle} \tag{3.2}$$

so that $(X, \|\cdot\|_{in})$ becomes a normed space and (X, d_{in}) becomes a metric space. Hence, we easily see that an inner product space is a normed space and thereby a metric space.

Remark. In IP3, the overbar denotes complex conjugation. Consequently, if X is a real vector space, we simply have

$$\langle x, y\rangle = \langle y, x\rangle \quad \text{(symmetry)}.$$

IP1 - IP3 imply

$$\left.\begin{aligned} &\text{a) } \langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle \\ &\text{b) } \langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle \end{aligned}\right\} \tag{3.3}$$

a) tells us that the inner product is linear in the first factor and conjugate linear in the second factor. Expressing them together, we say that the "inner product" is sesquilinear.

Now, if $x, y \in X$ and $(X, \langle\cdot,\cdot\rangle)$ is an inner product space, then the induced norm $\|\cdot\|_{in}$ on X satisfies the following

$$\begin{aligned} \|x - y\|_{in}^2 + \|x + y\|_{in}^2 &= \langle x - y, x - y\rangle + \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle y, y\rangle + \langle x, x\rangle + \langle y, y\rangle \\ &= 2(\|x\|_{in}^2 + \|y\|_{in}^2) \end{aligned} \tag{3.4}$$

The equality (3.4) is called the Parallelogram equality.

Definition 3.2 (Orthogonality)**.** An element x of an inner product space $(X, \langle\cdot,\cdot\rangle)$ is said to be orthogonal to an element $y \in X$ if $\langle x, y\rangle = 0$. We write it as $x \perp y$. Similarly, for subsets $A, B \subset X$ we write $x \perp A$ if $x \perp a$, $\forall a \in A$, and $A \perp B$ if $a \perp b$, $\forall a \in A$ and $\forall b \in B$.

Example 3.1 (Some examples of Inner Product Spaces).

1. Euclidean space $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_{Euc})$ is a Hilbert space with inner product defined by

$$\langle x, y\rangle = \xi_1\eta_1 + \cdots + \xi_n\eta_n \tag{3.5}$$

where $x = (\xi_j) = (\xi_1, \dots, \xi_n)$ and $y = (\eta_j) = (\eta_1, \dots, \eta_n)$ are vectors in $\mathbb{R}^n$.

2. Unitary space $(\mathbb{C}^n, \langle\cdot,\cdot\rangle_{\text{uni}})$ is a Hilbert space with inner product given by

$$\langle z, w\rangle = \xi_1\overline{\eta}_1 + \cdots + \xi_n\overline{\eta}_n \tag{3.6}$$

where $z = (\xi_j) = (\xi_1, \dots, \xi_n)$ and $w = (\eta_j) = (\eta_1, \dots, \eta_n)$ are vectors in $\mathbb{C}^n$.

3. Square integrable function space $L^2[a, b]$. The norm

$$\|x\|_2 = \left(\int_a^b |x(t)|^2 dt\right)^{1/2} \tag{3.7}$$

has been introduced before which can be induced from the inner product

$$\langle x, y\rangle_2 = \int_a^b x(t)\overline{y(t)}dt \tag{3.8}$$

It can be shown that $(L^2[a, b], \langle\cdot,\cdot\rangle_2)$ is complete with respect to the metric induced from the inner product (3.8). In other words, $L^2[a, b]$ is a Hilbert space.

4. The p-summable sequence space ℓ^p with $p \neq 2$ is not an inner product space, hence not a Hilbert space.

The last one is a quite illuminating and hence, its proof is given below.

Proof. The statement under tells us that the norm of ℓ^p with $p \neq 2$ can't be induced from an inner product. Had the norm been induced from the inner product, it would satisfy the parallelogram equality. Recall the ℓ^p norm

$$\|x\|_p = \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{1/p} \tag{3.9}$$

for $x = (\xi_j) \in \ell^p$. Now take $y = (1, 1, 0, 0, \dots) \in \ell^p$ and $z = (1, -1, 0, 0, \dots) \in \ell^p$. Then one has,

$$\|y\|_p = \|z\|_p = 2^{1/p} \tag{3.10}$$

And $y + z = (2, 0, 0, \dots)$ and $y - z = (0, 2, 0, \dots)$, so that

$$\begin{aligned} \|y + z\|_p &= (2^p)^{1/p} = 2 \\ \|y - z\|_p &= (2^p)^{1/p} = 2 \end{aligned} \tag{3.11}$$

(3.10) and (3.11) imply that

$$\|y + z\|_p^2 + \|y - z\|_p^2 = 4 + 4 = 8 \tag{3.12}$$

while

$$\begin{aligned} 2(\|y\|_p^2 + \|z\|_p^2) &= 2\left[(2^{1/p})^2 + (2^{1/p})^2\right] \\ &= 2 \cdot 2 \cdot 2^{2/p} \\ &= 4 \cdot 2^{2/p} \end{aligned} \tag{3.13}$$

From (3.12) and (3.13), one concludes that

$$\|y + z\|_p^2 + \|y - z\|_p^2 \neq 2(\|y\|_p^2 + \|z\|_p^2) \quad \text{if } p \neq 2$$

Hence, the ℓ^p norm $\|\cdot\|_p$ on the space of p-summable sequences is not induced from an inner product for $p \neq 2$. Hence, although $(\ell^p, \|\cdot\|_p)$ for $p \neq 2$ is a Banach space, it is not an inner product space and hence not a Hilbert space. ∎

§3.2 Further properties of inner product spaces

Lemma 3.1 (Schwarz inequality, triangle inequality)

An inner product $\langle \cdot, \cdot \rangle$ and the corresponding norm $\| \cdot \|_{\text{in}}$ associated with the inner product space $(X, \langle \cdot, \cdot \rangle)$ satisfy the Schwarz inequality and the triangle inequality as follows.

(a) (Schwarz inequality) We have

$$|\langle x, y \rangle| \leq \|x\|_{\text{in}} \|y\|_{\text{in}} \tag{3.14}$$

where the equality sign holds if and only if $\{x, y\}$ is a linearly dependent set.

(b) (Triangle inequality) The norm $\| \cdot \|_{\text{in}}$ also satisfies

$$\|x + y\|_{\text{in}} \leq \|x\|_{\text{in}} + \|y\|_{\text{in}} \tag{3.15}$$

where the equality sign holds if and only if $y = 0$ or $x = cy$ $(c > 0)$.

Proof. (a) If $y = 0$, then (3.14) automatically holds. Let $y \neq 0_x$. For every scalar α, one has

$$\begin{aligned} 0 \leq \|x - \alpha y\|_{\text{in}}^2 &= \langle x - \alpha y, x - \alpha y \rangle \\ &= \langle x, x \rangle - \langle x, \alpha y \rangle - \langle \alpha y, x \rangle + \langle \alpha y, \alpha y \rangle \\ &= \langle x, x \rangle - \overline{\alpha} \langle x, y \rangle - \alpha \langle y, x \rangle + \alpha \overline{\alpha} \langle y, y \rangle \\ &= \langle x, x \rangle - \overline{\alpha} \langle x, y \rangle - \alpha \overline{\langle x, y \rangle} + \alpha \overline{\alpha} \langle y, y \rangle \end{aligned}$$

Now, if we choose $\overline{\alpha} = \frac{\langle x,y \rangle}{\langle y,y \rangle}$, then the term in the square bracket on the right side above vanishes. The remaining inequality then is

$$\begin{aligned} 0 &\leq \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, x \rangle \\ &= \|x\|_{\text{in}}^2 - \frac{|\langle x, y \rangle|^2}{\|y\|_{\text{in}}^2} \\ \implies |\langle x, y \rangle|^2 &\leq \|x\|_{\text{in}}^2 \|y\|_{\text{in}}^2 \\ |\langle x, y \rangle| &\leq \|x\|_{\text{in}} \|y\|_{\text{in}} \end{aligned} \tag{3.16}$$

One, therefore, finds:

$$\begin{aligned} &0 \leq \|x\|_{\text{in}}^2 - \frac{|\langle x, y \rangle|^2}{\|y\|_{\text{in}}^2} \\ &\implies 0 \leq \|x\|_{\text{in}}^2 \|y\|_{\text{in}}^2 - |\langle x, y \rangle|^2 \quad \text{[By multiplying both sides of the above by } \|y\|_{\text{in}}^2\text{]} \\ &\implies |\langle x, y \rangle|^2 \leq \|x\|_{\text{in}}^2 \|y\|_{\text{in}}^2 \end{aligned}$$

Equality holds in this derivation if and only if $y = 0$ or $0 = \|x - \alpha y\|_{\text{in}}^2$, so that $x - \alpha y = 0$, i.e., $x = \alpha y$ or in other words, $\{x, y\}$ is a linearly dependent set in X. *(any set consisting of a zero vector is always linearly dependent)*

(b) One has

$$\begin{aligned} \|x + y\|_{\text{in}}^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \end{aligned} \tag{3.17}$$

Now, by the Cauchy-Schwarz inequality,

$$|\langle x, y \rangle| = |\langle y, x \rangle| \leq \|x\|_{\text{in}} \|y\|_{\text{in}} \tag{3.18}$$

By (3.17) and (3.18), one, therefore, obtains,

$$\begin{aligned}\|x+y\|_{\text{in}}^2 &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &\le \|x\|_{\text{in}}^2 + |\langle x,y\rangle| + |\langle y,x\rangle| + \|y\|_{\text{in}}^2 \quad \text{[By triangle inequality of numbers]} \\ &\le \|x\|_{\text{in}}^2 + 2\|x\|_{\text{in}}\|y\|_{\text{in}} + \|y\|_{\text{in}}^2 \\ &= (\|x\|_{\text{in}} + \|y\|_{\text{in}})^2 \end{aligned} \tag{3.19}$$

Taking square root on both sides of (3.19), one obtains,

$$\|x+y\|_{\text{in}} \le \|x\|_{\text{in}} + \|y\|_{\text{in}} \tag{3.20}$$

Equality holds in (3.20) if and only if

$$\langle x,y\rangle + \langle y,x\rangle = 2\|x\|_{\text{in}}\|y\|_{\text{in}} \tag{3.21}$$

But

$$\begin{aligned}\langle x,y\rangle + \langle y,x\rangle &= \langle x,y\rangle + \overline{\langle x,y\rangle} \\ &= 2\operatorname{Re}\langle x,y\rangle\end{aligned}$$

Hence, the condition (3.21) reduces to

$$\begin{aligned} 2\operatorname{Re}\langle x,y\rangle &= 2\|x\|_{\text{in}}\|y\|_{\text{in}} \\ \implies \operatorname{Re}\langle x,y\rangle &= \|x\|_{\text{in}}\|y\|_{\text{in}} \ge |\langle x,y\rangle| \quad \text{[By part (a)]}\end{aligned} \tag{3.22}$$

Since the real part of a complex number can't exceed the modulus of it, one must have an equality in (3.22):

$$\|x\|_{\text{in}}\|y\|_{\text{in}} = |\langle x,y\rangle| \tag{3.23}$$

Then, by part (a), one has $y = 0$ or $x = cy$.

∎

Lemma 3.2

If in an inner product space $(X, \langle\cdot,\cdot\rangle)$, one has $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.

Proof.

$$\begin{aligned}|\langle x_n,y_n\rangle - \langle x,y\rangle| &= |\langle x_n,y_n\rangle - \langle x_n,y\rangle + \langle x_n,y\rangle - \langle x,y\rangle| \\ &= |\langle x_n, y_n - y\rangle + \langle x_n - x, y\rangle| \\ &\le |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \quad \text{[Triangle inequality of numbers]} \\ &\le \|x_n\|_{\text{in}}\|y_n - y\|_{\text{in}} + \|x_n - x\|_{\text{in}}\|y\|_{\text{in}} \quad \text{[By Schwarz inequality]}\end{aligned} \tag{3.24}$$

Now, since $x_n \to x$ and $y_n \to y$ so that $\|x_n - x\|_{\text{in}} \to 0$ and $\|y_n - y\|_{\text{in}} \to 0$, one obtains from (3.24),

$$|\langle x_n,y_n\rangle - \langle x,y\rangle| \to 0$$

∎

Remark 3.1. Recall Theorem 2.28 in the context of the inner product map $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C}$ (considering X is a complex vector space). d_{in} is the metric on X induced by the inner product. Then Lemma 2 tells us that $\langle\cdot,\cdot\rangle : (X \times X, d_{\text{in}} \times d_{\text{in}}) \to (\mathbb{C}, d_{uni})$ is sequentially continuous at $(x,y) \in X \times X$. Indeed, one shows that for (x_n, y_n) converging to (x,y) in $(X \times X, d_{\text{in}} \times d_{\text{in}})$, one has $\langle\cdot,\cdot\rangle(x_n,y_n) = \langle x_n, y_n\rangle$ converging to $\langle\cdot,\cdot\rangle(x,y) = \langle x,y\rangle$ in $(\mathbb{C}, d_{uni})$. But in the context of metric spaces sequential continuity at $(x,y) \in X \times X$ is equivalent to continuity of a given map at that point. Hence, the inner product map $\langle\cdot,\cdot\rangle : (X \times X, d_{\text{in}} \times d_{\text{in}}) \to (\mathbb{C}, d_{uni})$ is continuous at a given $(x,y) \in X \times X$.

An isomorphism $T : (X, \langle\cdot,\cdot\rangle_X) \to (\tilde{X}, \langle\cdot,\cdot\rangle_{\tilde{X}})$ between inner product spaces (both X and $\tilde{X}$ are vector spaces over the same field) is a bijective linear operator which preserves the inner product, i.e., $\forall x, y \in X$,

$$\langle Tx, Ty\rangle_{\tilde{X}} = \langle x, y\rangle_X \tag{3.25}$$

In fact, by (3.25), one sees that

$$\begin{aligned}\langle Tx, Tx\rangle_{\tilde{X}} &= \langle x, x\rangle_X \\ \|Tx\|_{\tilde{X},\mathrm{in}} &= \|x\|_{X,\mathrm{in}}\end{aligned} \tag{3.26}$$

where $\|\cdot\|_{X,\mathrm{in}}$ is the norm induced from the inner product on X and $\|\cdot\|_{\tilde{X},\mathrm{in}}$ is the norm induced from the inner product on $\tilde{X}$. Then (3.26) tells us that

$$T : (X, \|\cdot\|_{X,\mathrm{in}}) \to (\tilde{X}, \|\cdot\|_{\tilde{X},\mathrm{in}})$$

is actually a bounded linear operator defined by (3.25) and hence continuous by Theorem 4.2.

§3.3 Orthonormality

Definition 3.3 (Orthonormal Set)**.** An orthonormal set M in an inner product space $(X, \langle\cdot,\cdot\rangle)$ is a subset $M \subseteq X$ whose elements are pairwise orthogonal. An orthonormal set is an orthogonal set whose elements have a unit norm, i.e., $\forall x, y \in M$

$$\langle x, y\rangle = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases} \tag{3.27}$$

If an orthogonal/orthonormal set M is countable, we can arrange it in a sequence (x_n) and call it orthogonal/orthonormal sequence.

More generally, an indexed family $(x_\alpha)_{\alpha\in I}$ is called orthogonal if $x_\alpha \perp x_\beta$ for all $\alpha, \beta \in I$, $\alpha \neq \beta$. The family $(x_\alpha)_{\alpha\in I}$ is called orthonormal if it is orthogonal and $\|x_\alpha\| = 1$, $\alpha \in I$, i.e.,

$$\langle x_\alpha, x_\beta\rangle = \delta_{\alpha\beta} \tag{3.28}$$

For orthogonal elements $x, y, x \neq y$ in the inner product space $(X, \langle\cdot,\cdot\rangle)$, one has $\langle x, y\rangle = 0$ so that

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x, x\rangle + \langle y, y\rangle && [\because \langle x, y\rangle = \langle y, x\rangle = 0] \\ &= \|x\|^2 + \|y\|^2\end{aligned} \tag{3.29}$$

More generally, if $\{x_1, \ldots, x_n\}$ is an orthogonal set, then

$$\|x_1 + \cdots + x_n\|^2_{\mathrm{in}} = \|x_1\|^2_{\mathrm{in}} + \|x_2\|^2_{\mathrm{in}} + \cdots + \|x_n\|^2_{\mathrm{in}} \tag{3.30}$$

Infact $\langle x_i, x_j\rangle = 0$ if $i \neq j$; consequently,

$$\left\|\sum_{j=1}^n x_j\right\|^2_{\mathrm{in}} = \left\langle \sum_{j=1}^n x_j, \sum_{k=1}^n x_k \right\rangle = \sum_{j=1}^n \langle x_j, x_j\rangle = \sum_{j=1}^n \|x\|^2_{\mathrm{in}}$$

Lemma 1 (Linear independence) An orthonormal set is linearly independent.

Proof. Let $\{e_1, \ldots, e_n\}$ be an orthonormal set in the inner product space $(X, \langle\cdot,\cdot\rangle)$. Consider the equation

$$\sum_{k=1}^{n} \alpha_k e_k = 0_x \tag{3.31}$$

Taking inner product with a fixed $e_j \in \{e_1, \ldots, e_n\}$ on both sides of (3.31) yields,

$$\begin{aligned} \left\langle \sum_{k=1}^{n} \alpha_k e_k, e_j \right\rangle &= \langle 0, e_j \rangle = 0 \\ \implies \sum_{k=1}^{n} \alpha_k \langle e_k, e_j \rangle &= 0 \\ \implies \alpha_j \langle e_j, e_j \rangle &= 0 \quad [\text{using orthogonality of } \{e_1, \ldots, e_n\}] \\ \implies \alpha_j &= 0 \quad \because \|e_j\|_{\text{in}} = 1] \end{aligned}$$

Since e_j was arbitrarily chosen from $\{e_1, \ldots, e_n\}$, one has $\alpha_j = 0$ for all $j \in \{1, \ldots, n\}$, proving linear independence for any finite orthonormal set.

If the orthonormal set that one starts with is infinite, by choosing any finite subcollection, one can prove that this finite subcollection is linearly independent as done above. In other words, one proves that the infinite orthonormal set is also linearly independent (any arbitrary subset of a vector space is linearly independent if every nonempty finite subset of it is linearly independent). ■

Example 3.2 (Unit Vectors of Euclidean Space)**.** Euclidean space $(\mathbb{R}^3, \langle \cdot, \cdot \rangle_{\text{Euc}})$. Here, the three unit vectors $\{(1,0,0), (0,1,0), (0,0,1)\}$ form an orthonormal set.

Example 3.3 (Schauder Basis of an Orthonormal Sequence)**.** The Schauder basis for the inner product space ℓ^2 is an orthonormal sequence (e_n), where $e_n = (\delta_{nj})$ is a sequence whose n-th element is 1 and all others are zero.

Example 3.4 (Prelude to Fourier Series)**.** Real-valued continuous functions $\mathcal{C}[0, 2\pi]$ on $[0, 2\pi]$ with inner product defined by

$$\langle x, y \rangle = \int_0^{2\pi} x(t) y(t) dt \tag{3.32}$$

is an inner product space, as earlier (although incomplete).

An orthogonal sequence in $(\mathcal{C}[0, 2\pi], \langle \cdot, \cdot \rangle)$ is (u_n), where

$$u_n(t) = \cos nt, \quad n = 0, 1, 2, \ldots$$

Another orthogonal sequence in $(\mathcal{C}[0, 2\pi], \langle \cdot, \cdot \rangle)$ is

$$u_n(t) = \sin nt, \quad n = 1, 2, \ldots$$

Explicit computation of inner product leads us to the following when $m \neq n$:

$$\begin{aligned} \langle u_m, u_n \rangle &= \frac{1}{2} \int_0^{2\pi} [\cos(m+n)t + \cos(m-n)t] dt \\ &= \frac{1}{2} \left[\frac{\sin(m+n)t}{m+n} + \frac{\sin(m-n)t}{m-n} \right]_0^{2\pi} \\ &= 0 \end{aligned} \tag{3.33}$$

when $m = n$, one has

$$\begin{aligned}\langle u_m, u_n\rangle &= \frac{1}{2}\int_0^{2\pi} \cos^2 mtdt \\ &= \frac{1}{2}\int_0^{2\pi} \left(\frac{1+\cos 2mt}{2}\right) dt \\ &= \pi + \frac{1}{4m}\left[\sin 2mt\right]_0^{2\pi} \\ &= \pi \end{aligned} \tag{3.34}$$

For $m = n = 0$, one has

$$\langle u_m, u_n\rangle = \int_0^{2\pi} 1dt = 2\pi t \tag{3.35}$$

Combining (3.33), (3.34) and (3.35), one has

$$\langle u_m, u_n\rangle = \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \neq 0 \\ 2\pi & \text{if } m = n = 0 \end{cases} \tag{3.36}$$

Similarly, one can go on to show that

$$\langle v_m, v_n\rangle = \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \neq 0, m, n = 1, 2, \ldots \end{cases} \tag{3.37}$$

Hence, from u_m, one can construct an orthonormal sequence (e_n) where

$$\left.\begin{aligned} e_0(t) &= \frac{1}{\sqrt{2\pi}}\cos 0t = \frac{1}{\sqrt{2\pi}}u_0 = \frac{1}{\sqrt{2\pi}} \\ e_n(t) &= \frac{1}{\sqrt{\pi}}\cos not = \frac{u_n(t)}{\|u_n\|} \quad n = 1, 2, \ldots \end{aligned}\right\} \tag{3.38}$$

From (v_n), one obtains the orthonormal sequence $(\tilde{e}_n)$ where

$$\tilde{e}_n(t) = \frac{1}{\sqrt{\pi}}\sin nt = \frac{v_n(t)}{\|v_n\|} \quad n = 1, 2, \ldots \tag{3.39}$$

When $m \neq n$

$$\begin{aligned}\langle v_m, v_n\rangle &= \frac{1}{2}\int_0^{2\pi} \sin mt \sin ntdt \\ &= \frac{1}{2}\int_0^{2\pi} \left[\cos(m-n)t - \cos(m+n)t\right] dt \\ &= \frac{1}{2}\left[\frac{\sin(m-n)t}{m-n} - \frac{\sin(m+n)t}{m+n}\right]_0^{2\pi} = 0 \end{aligned} \tag{3.40}$$

When $m = n$

$$\begin{aligned}\langle v_m, v_m\rangle &= \frac{1}{2}\int_0^{2\pi} \sin^2 mtdt \\ &= \frac{1}{2}\int_0^{2\pi} (1-\cos 2mt)\, dt \\ &= \frac{1}{2}(2\pi) - \frac{1}{2}\int_0^{2\pi} \cos 2mtdt \\ &= \pi - \frac{1}{4m}\left[\sin 2mt\right]_0^{2\pi} = \pi \end{aligned}$$

Exercise: Show that

$$u_m \perp u_n \quad \text{for } m \neq n$$

A great advantage of orthonormal sequences over arbitrary linearly independent sequences is the following:

If $(e_1, e_2, \ldots, e_n)$ is an orthonormal sequence in an inner product space $(X, \langle \cdot, \cdot \rangle)$ and we have $x \in \text{span}\{e_1, \ldots, e_n\}$ with n fixed, then one has

$$x = \sum_{k=1}^{n} \alpha_k e_k \tag{3.41}$$

for some scalars α_k. Now, take inner product of x with e_j for some fixed $j \in \{1, \ldots, n\}$ on both sides of (3.41) to obtain

$$\begin{aligned} \langle x, e_j \rangle &= \left\langle \sum_{k=1}^{n} \alpha_k e_k, e_j \right\rangle \\ &= \sum_{k=1}^{n} \alpha_k \langle e_k, e_j \rangle \\ &= \alpha_j \quad [\text{using orthonormality of } \{e_1, \ldots, e_n\}] \end{aligned} \tag{3.42}$$

With α_j determined by (3.41), (3.42) now takes the following form:

$$x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k \tag{3.43}$$

This shows that determination of the unknown coefficients α_k's is simpler for $x \in \text{span}\{e_1, \ldots, e_n\}$ given by (3.41). Another usefulness for orthonormality becomes apparent if in (3.41) or in (3.43), we want to add another term $\alpha_{n+1} e_{n+1}$ to take care of an $x \in \text{span}\{e_1, \ldots, e_{n+1}\}$. In this case, one just needs to compute α_{n+1} as $\alpha_1, \ldots, \alpha_n$ remain unchanged.

More generally, if we consider $x \in X$, not necessarily contained in $Y_n = \text{span}\{e_1, \ldots, e_n\}$, one can define a new $y \in Y_n$ by setting

$$y = \sum_{k=1}^{n} \langle x, e_k \rangle_X e_k \tag{3.44}$$

and then define $z \in X$ by setting

$$x = y + z \tag{3.45}$$

i.e.,

$$z = x - y \tag{3.46}$$

We, first of all, want to show that z given by (3.46) satisfies $z \perp y$.

First, note that by orthonormality, one has

$$
\begin{aligned}
\|y\|_X^2 &= \langle y, y\rangle \\
&= \left\langle \sum_{k=1}^{n} \langle x, e_k\rangle e_k, \sum_{m=1}^{n} \langle x, e_m\rangle e_m \right\rangle_X \\
&= \sum_{k=1}^{n} \sum_{m=1}^{n} \langle x, e_k\rangle_X \overline{\langle x, e_m\rangle_X} \langle e_k, e_m\rangle_X \\
&= \sum_{k=1}^{n} \sum_{m=1}^{n} \langle x, e_k\rangle_X \overline{\langle x, e_m\rangle_X} \delta_{km} \\
&= \sum_{k=1}^{n} \langle x, e_k\rangle_X \overline{\langle x, e_k\rangle_X} \\
&= \sum_{k=1}^{n} |\langle x, e_k\rangle_X|^2 \qquad (3.47)
\end{aligned}
$$

Now,

$$
\begin{aligned}
\langle z, y\rangle_X = \langle x - y, y\rangle_X &= \langle x, y\rangle_X - \|y\|_X^2 \\
&= \left\langle x, \sum_{k=1}^{n} \langle x, e_k\rangle_X e_k \right\rangle_X - \|y\|_{\text{in}}^2 \\
&= \sum_{k=1}^{n} \langle x, e_k\rangle_X \overline{\langle x, e_k\rangle_X} - \|y\|_{\text{in}}^2 \\
&= 0 \quad [\text{by } (3.47)]
\end{aligned}
$$

Now, since $z \perp y$ by Pythagorean relation on (3.45),

$$
\begin{aligned}
\|z\|_{\text{in}}^2 &= \|y + z\|_{\text{in}}^2 = \|y\|_{\text{in}}^2 + \|z\|_{\text{in}}^2 \\
\|z\|_{\text{in}}^2 &= \|x\|_{\text{in}}^2 - \|y\|_{\text{in}}^2 \\
&= \|x\|_{\text{in}}^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle_X|^2 \qquad (3.48)
\end{aligned}
$$

Since $\|z\|_{\text{in}}^2 \geq 0$ for every $n = 1, 2, \ldots,$ one has from (3.48),

$$
\|x\|_{\text{in}}^2 \geq \sum_{k=1}^{n} |\langle x, e_k\rangle_X|^2 \tag{3.49}
$$

Now, $(\sum_{k=1}^{n} |\langle x, e_k\rangle_X|^2)$ is a monotone non-decreasing sequence that is bounded by $\|x\|_X^2$ given in (3.49). Hence, the sequence of the partial sum of the series $\sum_{k=1}^{\infty} |\langle x, e_k\rangle_X|^2$ converges and

$$
\sup \left(\sum_{k=1}^{n} |\langle x, e_k\rangle_X|^2 \right)_{n=1}^{\infty} = \sum_{k=1}^{\infty} |\langle x, e_k\rangle_X|^2 \leq \|x\|_X^2 \tag{3.50}
$$

which is formally known as Bessel inequality.

Theorem 3.3 (Bessel Inequality)

Let (e_k) be an orthonormal sequence in an inner product space $(X, \langle \cdot, \cdot \rangle_X)$. Then for every $x \in X$, one has

$$
\sum_{k=1}^{\infty} |\langle x, e_k\rangle_X|^2 \leq \|x\|_X^2 \quad \text{(Bessel inequality)} \tag{3.51}
$$

The inner products $\langle x, e_k\rangle_X$ in (3.51) are called **Fourier coefficients** of $x \in X$ with the orthonormal sequence (e_k).

• **How to obtain an orthonormal sequence from an arbitrary linearly independent sequence in an inner product space:** This is accomplished by a constructive procedure, called the **Gram-Schmidt orthonormalization process**.

Let (x_i) be a given linearly independent sequence in an inner product space $(X, \langle\cdot,\cdot\rangle_X)$. The resulting orthonormal sequence (e_i) will have the property that for every $n \in \mathbb{N}$,

$$\text{span}\{e_1, \ldots, e_n\} = \text{span}\{x_1, \ldots, x_n\} \tag{3.52}$$

The process is as follows:

Step 1. The first element of the sequence (e_n) is

$$e_1 = \frac{x_1}{\|x_1\|_{\text{in}}} \tag{3.53}$$

2nd Step

Write x_2 as:

$$x_2 = \langle x_2, e_1\rangle_X e_1 + v_2 \tag{3.54}$$

So that

$$v_2 = x_2 - \langle x_2, e_1\rangle_X e_1 \tag{3.55}$$

Now, (3.55) tells us that

$$v_2 = x_2 - \left\langle x_2, \frac{x_1}{\|x_{\text{in}}\|_X}\right\rangle_X \frac{x_1}{\|x_1\|_{\text{in}}} \tag{3.56}$$

The right side of (3.56) is a nonzero linear combination of x_1 and x_2, i.e., the coefficients of the linear combination above are not all zero. Hence, by linear independence of $\{x_1, x_2\}$, the right side of (27), i.e., v_2 can't be zero, so that we can take

$$e_2 = \frac{v_2}{\|v_2\|_X} \tag{3.57}$$

Of course, $\|e_2\|_{\text{in}} = 1$, and

$$\begin{aligned}
\langle e_2, e_1\rangle &= \frac{1}{\|v_2\|_{\text{in}}}\langle v_2, e_1\rangle_X \\
&= \frac{1}{\|v_2\|_{\text{in}}}\langle x_2 - \langle x_2, e_1\rangle_X e_1, e_1\rangle_X \\
&= \frac{1}{\|v_2\|_{\text{in}}}\left[\langle x_2, e_1\rangle_X - \langle x_2, e_1\rangle_X \langle e_1, e_1\rangle_X\right] \\
&= \frac{1}{\|v_2\|_{\text{in}}}\left[\langle x_2, e_1\rangle_X - \langle x_2, e_1\rangle_X\right] \ \left[\because \|e_1\|_{\text{in}}^2 = 1 \text{ by construction in (3.53)}\right] \\
&= 0
\end{aligned}$$

3rd Step Construct the vector

$$v_3 = x_3 - \langle x_3, e_1\rangle_X e_1 - \langle x_3, e_2\rangle_X e_2 \tag{3.58}$$

Then $v_3 \neq 0_X$ follows from linear independence of $\{x_j\}$ and that $v_3 \perp e_1$ and $v_3 \perp e_2$ (verify!). Then

$$e_3 = \frac{v_3}{\|v_3\|_{\text{in}}} \tag{3.59}$$

4th Step

Construct the vector

$$v_n = x_n - \langle x_n, e_1\rangle_X e_1 - \cdots - \langle x_n, e_{n-1}\rangle_X e_{n-1}$$

using the orthonormal set $\{e_1, \ldots, e_{n-1}\}$ obtained in the previous $(n-1)$ steps. Rewrite the above compactly as,

$$v_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle_X e_k \tag{3.60}$$

It is immediately seen that $v_n \neq 0_X$ and v_n is orthogonal to e_k for all k constructed up to $(n-1)$ steps. Then, we take

$$e_n = \frac{v_n}{\|v_n\|_X} \tag{3.61}$$

Series Related to Orthonormal Sequences

Example 3.5 (Fourier series)**.** A **trigonometric series** is a series of the form

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt) \tag{3.62}$$

A real-valued function x on $\mathbb{R}$ is said to be **periodic** if there is a positive number p (called a **period** of x) such that

$$x(t+p) = x(t) \tag{3.63}$$

Now, let x be of period 2π and continuous. By definition, the **Fourier series** of x is the trigonometric series (3.62) with coefficients a_k and b_k given by the **Euler formula**:

$$\left.\begin{aligned} a_0 &= \frac{1}{2\pi} \int_0^{2\pi} x(t)dt \\ a_k &= \frac{1}{\pi} \int_0^{2\pi} x(t) \cos ktdt, \quad k = 1, 2, \ldots \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} x(t) \sin ktdt, \quad k = 1, 2, \ldots \end{aligned}\right\} \tag{3.64}$$

The coefficients a_k and b_k are called **Fourier coefficients** of x. If the Fourier series of x converges for each t and has the sum $x(t)$, then we write

$$x(t) = a_0 + \sum_{k=0}^{\infty} (a_k \cos kt + b_k \sin kt) \tag{3.65}$$

Since x is periodic of period 2π, in (3.64), we may replace the interval of integration $[0, 2\pi]$ by any other interval of length 2π, for instance $[-\pi, \pi]$.

Now, consider the orthogonal sequences $(u_k)_{k=0}^{\infty}$ and $(v_k)_{k=1}^{\infty}$ studied in example 7.3. Using these orthogonal sequences, (3.65) can be rewritten as

$$x(t) = a_0 u_0(t) + \sum_{k=1}^{\infty} [a_k u_k(t) + b_k v_k(t)] \tag{3.66}$$

Using the orthogonality of (u_n) and (v_n) and the fact that $\langle u_j, u_k \rangle = \langle v_j, v_k \rangle = 0$ if $j \neq k$,

$$\begin{aligned} \langle x, u_j \rangle &= a_0 \langle u_0, u_j \rangle + \sum_{k=1}^{\infty} [a_k \langle u_k, u_j \rangle + b_k \langle v_k, u_j \rangle] \\ &= a_j \langle u_j, u_j \rangle \\ &= a_j \|u_j\|^2 = \begin{cases} 2\pi a_0 & \text{if } j = 0 \\ \pi a_j & \text{if } j = 1, 2, \ldots \end{cases} \quad \text{[follows from (3.33) and (3.34)]} \end{aligned} \tag{3.67}$$

Similarly, if we multiply (3.66) with $v_j(t)$ on both sides and integrate over $[0, 2\pi]$ and apply orthogonality of (u_k) and (v_k) as well as $u_j \perp v_j, \forall j, k$,

$$\begin{aligned}\langle x, v_j\rangle &= \sum_{k=1}^{\infty}\Big[a_k\langle u_j, v_k\rangle + b_k\langle v_k, v_j\rangle\Big] \\ &= b_j\|v_j\|^2 && (3.68)\\ &= \pi b_j \quad \text{[By the fact that } \langle u_i, v_j\rangle = \pi \text{ when } m = n \text{ as was shown earlier]} && (3.69)\end{aligned}$$

Here, (3.67) and (3.69) hold true for $j = 0, 1, 2, \ldots$. Now, recall the orthonormal sequences (e_j) and $(\tilde{e}_j)$ as constructed in Example 7.3:

$$e_0 = \frac{u_0}{\|u_0\|} \quad \text{and} \quad e_j = \frac{u_j}{\|u_j\|} \quad j = 1, 2, \ldots \tag{3.70}$$

and

$$\tilde{e}_j = \frac{v_j}{\|v_j\|} \quad j = 1, 2, \ldots \tag{3.71}$$

Using (3.70) in (3.67) and (3.69), one obtains,

$$\left.\begin{aligned} a_j &= \frac{\langle x, e_j\rangle}{\|u_j\|} \quad j = 0, 1, 2, \ldots \\ b_j &= \frac{\langle x, \tilde{e}_j\rangle}{\|\tilde{v}_j\|} \quad j = 1, 2, \ldots \end{aligned}\right\} \tag{3.72}$$

Therefore, in (3.72), one has

$$a_k u_k(t) = \frac{1}{\|u_k\|}\langle x, e_k\rangle u_k(t) = \langle x, e_k\rangle e_k(t) \quad \text{[By (3.69) and (3.70)]} \tag{3.73}$$

for $k = 0, 1, 2, \ldots$ and

$$b_k v_k(t) = \frac{1}{\|v_k\|}\langle x, \tilde{e}_k\rangle v_k(t) = \langle x, \tilde{e}_k\rangle \tilde{e}_k(t) \quad \text{[Again by (3.69) and (3.70)]} \tag{3.74}$$

for $k = 1, 2, \ldots$
Hence, using these orthonormal sequences, (3.66) can be written as

$$x = \langle x, e_0\rangle e_0 + \sum_{k=1}^{\infty}[\langle x, e_k\rangle e_k + \langle x, \tilde{e}_k\rangle \tilde{e}_k] \tag{3.75}$$

The Fourier series expansion of x given by (3.75) explains why the inner products $\langle x, e_k\rangle$ are called Fourier coefficients w.r.t. the orthonormal sequence $\{x_k\}$.

Convergence of Fourier Series: Given any orthonormal sequence (e_n) in a Hilbert space $\mathcal{H}$, we may consider infinite series of the form

$$\sum_{k=1}^{\infty} \alpha_k e_k \tag{3.76}$$

where $\alpha_1, \alpha_2, \ldots$ are any scalars. We say that the series (3.76) converges and has the sum

$$S = \sum_{k=1}^{\infty} \alpha_k e_k$$

if there exists an $S \in \mathcal{H}$ such that the sequence (S_n) of the partial sums

$$S_n = \alpha_1 e_1 + \cdots + \alpha_n e_n \tag{3.77}$$

converges to S, i.e., $\|S_n - S\| \to 0$ as $n \to \infty$.

Theorem 3.4 (Convergence of Fourier Series)
Let (e_k) be an orthonormal sequence in a Hilbert space $\mathcal{H}$. Then:

(a) The series (45) converges (in the norm on $\mathcal{H}$) if and only if the following series converges:

$$\sum_{k=1}^{\infty} |\alpha_k|^2 \tag{3.78}$$

(b) If the series (3.76) converges, then the coefficients α_k are the Fourier coefficients $\langle x, e_k \rangle$ where

$$x = \sum_{k=1}^{\infty} \alpha_k e_k \tag{3.79}$$

Hence, in this case the finite sum (3.76) can be written as

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \tag{3.80}$$

with $x_k = \langle x, e_k \rangle$.

(c) Convergence of (b): For any $x \in \mathcal{H}$, the series (3.76) with $\alpha_k = \langle x, e_k \rangle$ converges (in the norm of $\mathcal{H}$).

Proof. (a) Let $S_n = \alpha_1 e_1 + \cdots + \alpha_n e_n$ and $\sigma_n = |\alpha_1|^2 + \cdots + |\alpha_n|^2$. Then due to orthonormality for any m and $n > m$,

$$\begin{aligned}
\|S_n - S_m\|^2 &= \|\alpha_{m+1} e_{m+1} + \cdots + \alpha_n e_n\|^2 \\
&= |\alpha_{m+1}|^2 \|e_{m+1}\|^2 + \cdots + |\alpha_n|^2 \|e_n\|^2 \\
&= |\alpha_{m+1}|^2 + \cdots + |\alpha_n|^2 \\
&= \sigma_n - \sigma_m
\end{aligned} \tag{3.81}$$

From (3.81), it's clear that (S_n) is Cauchy in $\mathcal{H}$ if and only if (σ_n) is Cauchy in $\mathbb{R}$. Since both H and $\mathbb{R}$ are complete, it follows that the sequence (S_n) of the partial sums converges if and only if the sequence (σ_n) converges. In other words, $\sum_{k=1}^{\infty} \alpha_k e_k$ converges if and only if $\sum_{k=1}^{\infty} |\alpha_k|^2$ converges.

(b) Let $\sum_{k=1}^{\infty} \alpha_k e_k$ converges. Set

$$x = \sum_{k=1}^{\infty} \alpha_k e_k$$

We know $\sum_{k=1}^{\infty} \alpha_k e_k$ is the limit of the sequence (S_n) of partial sums, where

$$S_n = \sum_{k=1}^{n} \alpha_k e_k = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

And

$$\langle S_n, e_j \rangle = \alpha_j, \quad \text{for } j = 1, \ldots, k \quad (k \text{ fixed}) \tag{3.82}$$

By assumption, $S_n \to x$, i.e., $\lim_{n\to\infty} S_n = x$. Then (k fixed)

$$\begin{aligned}
\lim_{n\to\infty} \langle S_n, e_j \rangle &= \langle \lim_{n\to\infty} S_n, e_j \rangle \quad \text{[From the continuity of the inner product]} \\
&= \langle x, e_j \rangle
\end{aligned}$$

i.e., $\lim_{n\to\infty} \alpha_j = \langle x, e_j \rangle$

$$\implies \alpha_j = \langle x, e_j \rangle \quad (j \leq k)$$

The crux of the matter is that one can take $k(n)$ as large as one pleases because $n \to \infty$. Hence, k can be taken as large as possible and one has

$$\alpha_j = \langle x, e_j \rangle, \quad \text{for every } j = 1, 2, \dots \tag{3.83}$$

(c) From Bessel inequality, for $x \in \mathcal{H}$, one knows that

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2,$$

which tells us that the series $\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$ is convergent. Then by part (a), one has

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

to be convergent, as well.

■

Lemma 3.5

Any x in an inner product space $(x, \langle \cdot, \cdot \rangle_x)$ can have at most countably many non-zero Fourier coefficients $\langle x, e_{\mathcal{K}} \rangle$ with respect to an orthonormal family (possibly uncountable) $(e_{\mathcal{K}})_{\mathcal{K} \in I}$ in X.

Remark 3.2. Lemma 3.5 will enable us to write (3.80) even when we have an uncountable orthonormal sequence $(e_{\mathcal{K}})_{\mathcal{K} \in I}$ in X.

4 Fundamental Theorems of Normed Spaces

§4.1 Zorn's Lemma

Definition 4.1. A partial ordering on a partially ordered set M is a binary relation $\leq$ that satisfies the following conditions:

(P01) $a \leq a$ for all $a \in M$ (reflexivity)

(P02) If $a \leq b$ and $b \leq a$, then $a = b$ (antisymmetry)

(P03) If $a \leq b$ and $b \leq c$, then $a \leq c$ (transitivity)

Remark 4.1. (P02) means that two unequal or distinct elements a and b can't both satisfy $a \leq b$ and $b \leq a$. Note that there can be elements in a partially ordered set that are not comparable. To phrase differently, M can contain elements a and b for which neither $a \leq b$ nor $b \leq a$ holds. In such situation, we say that a and b are *incomparable*. On the contrary, 2 element a and b of M are called *comparable elements* if they satisfy $a \leq b$ or $b \leq a$ (or both, which are the same thing according to antisymmetry).

Definition 4.2 (Totally ordered set or Chain)**.** A **Totally ordered set** or **Chain** is a partially ordered set such that every 2 elements of the set are comparable. In other words, a chain is a partially ordered set that has no incomparable elements.

Definition 4.3 (Upper bound)**.** An **upper bound** of a subset W of a partially ordered set M is an element $u \in M$ such that

$$x \leq u \quad \text{for all} \quad x \in W \tag{4.1}$$

Definition 4.4 (Maximal element)**.** A **maximal element** of a partially ordered set M is an element $m \in M$ such that

$$m \leq x \implies m = x \tag{4.2}$$

i.e, there is no x in M different from m satisfying $m \leq x$. In other words, if x and m are distinct in M, then either $x \leq m$ or x and m are incomparable.

Axiom 1 (Zorn's Lemma)**.** Let M be a partially ordered set in which every chain has an upper bound. Then M has a maximal element.

Theorem 4.1 (Hamel Basis)

Every vector space $X \neq \{0\}$ has a Hamel basis.

Proof. We take M to be the set of all linearly independent subsets of X. Since $X \neq \{0\}$, there exists $x \neq 0$ such that $x \in X$ so that $\{x\} \in M$ and hence $M \neq \varnothing$. Set inclusion defines a partial ordering on M. Note that there may be pair of elements of M that are incomparable under set inclusion. Now, every chain $C \subset M$ has an upper bound, namely, the union of elements of C. By zorn's lemma M has a maximal element that we denote by B. We now show that B is a Hamel basis of X.

Let $Y = \text{span } B$. Then Y is a subspace of X. In fact $Y = X$. Suppose the contrary, i.e, Y is a proper subspace of X so that there exists a $z \in X$ such that $z \notin Y = \text{span } B$. Since z is not in span B, $B \cup \{z\}$ is linearly independent subset of X so that $\{z\} \cup B$ belongs to M. One, therefore, has $B \subset \{z\} \cup B$ with $B \neq \{z\} \cup B$. This is a contradiction to the maximality of B. Hence span $B = X$, proving that B is a Hamel basis of X. ■

Theorem 4.2 (Total Orthonormal Set)

In every Hilbert space $H \neq \{0\}$, there exists a total orthonormal set.

Proof. Let M be the set of all orthonormal subsets of H. Since $H \neq \{0\}$, there exists $x \in H$ with $x \neq 0$. Let $y = \frac{x}{\|x\|}$. Then $\{y\} \in M$ so that $M \neq \varnothing$. Set inclusion defines a partial ordering on M. Every chain $C \subset M$ has an upper bound, namely, the union of all elements of C. Then by Zorn's lemma, M has a maximal element that we denote by F. We prove that F is total in H.

Assume for contradiction that F is not total in H. Since H is complete, we have a necessary and sufficient condition of totality given by (Theorem 1 of lecture 8 TODO reference). Since F is not total in H, there exists a noonzero vector $z \in H$ such that $z \perp F$. Take $e = \frac{z}{\|z\|}$ so that $F_1 = F \cup \{e\}$ is an orthonormal set in H. Since $F \subseteq F_1$ with $F \neq F_1$, F is not maximal, which is a contradiction. Hence F is total in H. ■

§4.2 Hahn-Banach Theorem

Definition 4.5 (Real Sublinear Functional). Let X be a real vector space. A function $p : X \to \mathbb{R}$ is called a **sublinear functional** if it satisfies the following properties:

1. $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$.
2. $p(\alpha x) = \alpha p(x)$ for all $x \in X$ and $\alpha \geq 0$.

Theorem 4.3 (Hahn-Banach Theorem for Real Vector Spaces)

Let X be a real vector space and p be a sublinear functional. Let $f : Z \to \mathbb{R}$ be a linear functional defined on subspace Z of X and satisfies,

$$f(x) \leq p(x) \quad \forall x \in Z.$$

Then, f has a linear extension $\tilde{f} : X \to \mathbb{R}$ such that,

$$\tilde{f}(x) \leq p(x) \quad \forall x \in X.$$

Proof. Let E be the set of all linear extensions g of f satisfying $g(x) \leq p(x)$ for all $x \in \mathcal{D}(g)$ Since $f \in E$, E is non-empty. Define a partial order on E by $g_1 \leq g_2$ iff $\mathcal{D}(g_1) \subset \mathcal{D}(g_2)$ and $g_2(x) = g_1(x)$ for all $x \in \mathcal{D}(g_1)$. Let $C \subset E$ be any chain in E.

Define $\hat{g} : \bigcup_{g \in C} \mathcal{D}(g) \to \mathbb{R}$ by $\hat{g}(x) = g(x)$ for all $x \in \mathcal{D}(g)$ for some $g \in C$ If $g_1, g_2 \in C$ such that $\mathcal{D}(g_1) \cap \mathcal{D}(g_2) \neq \varnothing$. Since C is totally ordered, either $g_1 \leq g_2$ or $g_2 \leq g_1$. In any of these cases, $g_1(x) = g_2(x) \forall x \in \mathcal{D}(g_1) \cap \mathcal{D}(g_2)$. Thus, $\hat{g}$ is well-defined. The domain of $\hat{g}$ is a vector space since C is a chain. Let $x, y \in \mathcal{D}(\hat{g})$ and α be any scalar. Then,

$$\begin{aligned} \hat{g}(\alpha x + y) &= g(\alpha x + y) \quad \text{for some } g \in C \\ &= \alpha g(x) + g(y) \\ &= \alpha \hat{g}(x) + \hat{g}(y). \end{aligned}$$

Therefore, $\hat{g}$ is a linear functional such that $g \leq \hat{g} \forall g \in C$. Therefore C has an upper bound in E. Since, C was arbitrary chain, by Zorn's Lemma, E has a maximal element $\tilde{f}$. By definition of E, $\tilde{f}$ is a linear extension of f and $\tilde{f}(x) \leq p(x), \forall x \in \mathcal{D}(\hat{f})$.

We want to show that $\mathcal{D}(\hat{f}) = X$. Suppose the contrary. Then, $\mathcal{D}(\hat{f}$ is a proper subspace of X. Let $y_1 \in X - \mathcal{D}(\hat{f})$. $y_1 \neq 0$ since $0 \in \mathcal{D}(\hat{f})$. Let $Y_1 = \text{span}\{y_1, \mathcal{D}(\hat{f})\}$. Let $x \in Y_1$. Then, x can be uniquely written as,

$$x = y + \alpha y_1 \quad \text{for some } y \in \mathcal{D}(\hat{f}) \text{ and } \alpha \in \mathbb{R}.$$

Define $g : Y_1 \to \mathbb{R}$ by

$$g(y + \alpha y_1) = \hat{f}(y) + \alpha c \tag{4.3}$$

For some real constant c. Since $\hat{f}$ is linear and αc can also be seen as a linear functional, g is a also linear. For all $y \in \mathcal{D}(\hat{f})$, $g(y + 0 \cdot y_1) = \hat{f}(y) = g(y)$. Therefore, g is a linear extension of $\hat{f}$ on Y_1.

Let $y, z \in \mathcal{D}(\hat{f})$. We have,

$$\begin{aligned}
\hat{f}(y) - \hat{f}(z) = \hat{f}(y - z) &\leq p(y - z) \\
&= p(y + y_1 - y_1 - z) \\
&\leq p(y + y_1) + p(-y_1 - z)
\end{aligned}$$

Hence, we have,

$$-p(-y_1 - z) - \hat{f}(z) \leq p(y + y_1) - \hat{f}(y) \tag{4.4}$$

Left hand side of (4.4) does not depend on y and right hand side does not depend on z. Let m_1 be the supremum of Left Hand side over z and m_2 be the infimum of Right Hand side over y. Then, we have,

$$-p(-y_1 - z) - \hat{f}(z) \leq c \quad \forall z \in \mathcal{D}(\hat{f}) \tag{4.5}$$

$$p(y + y_1) - \hat{f}(y) \geq c \quad \forall y \in \mathcal{D}(\hat{f}) \tag{4.6}$$

where $m_1 \leq c \leq m_2$. Let g be a linear extension of $\hat{f}$ defined in (4.3) using this particular c. We now show that $g(x) \leq p(x)$ for all $x \in Y_1$. Let $x = y + \alpha y_1 \in Y_1$ where $\alpha \in \mathbb{R}$.

Case 1: $\alpha < 0$. Let $z = \frac{y}{\alpha}$. Then, from equation (4.5), we have,

$$-p(-y_1 - \frac{y}{\alpha}) - \hat{f}(\frac{y}{\alpha}) \leq c$$

Multiplying by $-\alpha > 0$ on both sides, we get,

$$\begin{aligned}
&\alpha p(-y_1 - \frac{y}{\alpha}) - \hat{f}(y) \leq -\alpha c \\
\implies &\hat{f}(y) + \alpha c \leq -\alpha p(-y_1 - \frac{y}{\alpha}) \\
\implies &g(y + \alpha y_1) \leq p(y + \alpha y_1) \\
\implies &g(x) \leq p(x).
\end{aligned}$$

Case 2: $\alpha = 0$. Then $x = y$.

$$g(x) = g(y) = \hat{f}(y) \leq p(y) = p(x).$$

Case 3: $\alpha > 0$. Let $y = \frac{y}{\alpha}$. Then, from equation (4.6), we have,

$$c \le p\left(\frac{y}{\alpha} + y_1\right) - \hat{f}\left(\frac{y}{\alpha}\right)$$

Multiplying by $\alpha > 0$ on both sides, we get

$$\begin{aligned} \alpha c &\le \alpha p\left(\frac{y}{\alpha} + y_1\right) - \hat{f}(y) \\ \implies \hat{f}(y) + \alpha c &\le \alpha p\left(\frac{y}{\alpha} + y_1\right) \\ \implies g(y + \alpha y_1) &\le p(y + \alpha y_1) \\ \implies g(x) &\le p(x). \end{aligned}$$

Therefore, $g(x) \le p(x)$ for all $x \in Y_1$. Hence, $g \in E$ and $\hat{f} \le g$. This contradicts the maximality of $\hat{f}$. Therefore, $\mathcal{D}(\hat{f}) = X$. Hence $\hat{f}$ is the desired linear extension of f. ■

Definition 4.6 (Sublinear Functional)**.** Let X be a real or a complex vector space. A function $p : X \to \mathbb{R}$ is called a **sublinear functional** if it satisfies the following properties:

$$\begin{aligned} &p(x+y) \le p(x) + p(y) \quad \forall x, y \in X \\ &p(\alpha x) = |\alpha| p(x) \quad \forall x \in X \text{ and } \alpha \in \mathbb{R}. \end{aligned}$$

Theorem 4.4 (Generalized Hann-Banach Theorem)

Let X be a real or complex vector space and p be sublinear functional on X as defined in 4.6. Let $f : Z \to \mathbb{K}$ be a linear functional defined on subspace Z of X and satisfies,

$$|f(x)| \le p(x) \quad \forall x \in Z.$$

There there exists a linear extension $\tilde{f} : X \to \mathbb{K}$ of f such that,

$$|\hat{f}(x)| \le p(x) \quad \forall x \in X.$$

Proof. Let X be a real vector space. Then $|f(x)| \le p(x) \implies f(x) \le p(x)$ for all $x \in Z$. By Theorem 4.3, there exists a linear extension $\hat{f} : X \to \mathbb{R}$ of f such that $\hat{f}(x) \le p(x)$ for all $x \in X$.

$$-\hat{f}(x) = \hat{f}(-x) \le p(-x) = p(x) \quad \forall x \in X.$$

Hence, $|\hat{f}(x)| \le p(x)$ for all $x \in X$.

Let X be a complex vector space. Then Z is also a complex vector space. Hence $f(x)$ can be written as,

$$f(x) = f_1(x) + i f_2(x) \quad \text{where } f_1, f_2 : Z \to \mathbb{R}.$$

Let X_r and Z_r be real vector spaces by restricting scalar multiplication of X and Z to real scalars. Hence, X and X_r are same as sets and addition, but different only in scalar multiplication, and same goes for Z and Z_r. Since, f is linear in Z, f_1 and f_2 are real valued linear functionals on Z_r. Also, $f_1(x) \le |f(x)|$. Hence, we have

$$f_1(x) \le p(x) \quad \forall x \in Z_r.$$

Therefore, by Hann-Banach theorem of Real vector space, there exists a linear extension $\hat{f}_1$ of f such that,

$$\hat{f}_1(x) \le p(x) \quad \forall x \in X_r.$$

For every $x \in Z$, we have,

$$\begin{aligned} i(f_1(x) + if_2(x)) &= if(x) = f(ix) = f_1(ix) + if_2(ix) \\ \implies if_1(x) - f_2(x) &= f_1(ix) + if_2(ix) \end{aligned}$$

Taking real part on both sides, we get $f_1(ix) = -f_2(x)$ for all $x \in Z$. Hence, we can define $\hat{f} : X \to \mathbb{C}$ by

$$\hat{f}(x) = \hat{f}_1(x) - i\hat{f}_1(ix) \quad \forall x \in X.$$

So that, $\hat{f}(x) = f(x)$ for all $x \in Z$. Since, $\hat{f}_1$ is linear, $\hat{f}$ is linear on additivity, that is $\hat{f}(x+y) = \hat{f}(x) + \hat{f}(y)$ for all $x, y \in X$. Let $a + ib$ be any scalar. Then,

$$\begin{aligned} \hat{f}((a+ib)x) &= \hat{f}_1((a+ib)x) - i\hat{f}_1((a+ib)ix) \\ &= (a+ib)\hat{f}_1(x) - i(a+ib)\hat{f}_1(ix) \\ &= (a+ib)(\hat{f}_1(x) - i\hat{f}_1(ix)) \\ &= (a+ib)\hat{f}(x). \end{aligned}$$

Therefore, $\hat{f}$ is linear. Hence $\hat{f}$ is a linear extension of f.

If $\hat{f}(x) = 0$, then $\hat{f}(x) = 0 = |\hat{f}_1(0)| \leq p(0)$. Let $x \in X$ such that $\hat{f}(x) \neq 0$. Then, writting in polar form, we have,

$$\begin{aligned} \hat{f}(x) &= |\hat{f}(x)|e^{i\theta} \\ \implies |\hat{f}(x)| &= e^{-i\theta}\hat{f}(x) \end{aligned}$$

Hence, we have,

$$\begin{aligned} |\hat{f}(x)| &= e^{-i\theta}\hat{f}(x) \\ &= \hat{f}(e^{-i\theta}x) \\ &\leq p(e^{-i\theta}x) \\ &= |e^{-i\theta}|p(x) \\ &= p(x). \end{aligned}$$

Therefore, $|\hat{f}(x)| \leq p(x)$ for all $x \in X$. Hence $\hat{f}$ is the desired linear extension of f. ■

Theorem 4.5 (Hann-Banach Theorem for Normed Spaces)
Let f be a bounded linear functional on a subspace Z of a normed space X. Then there exists a bounded linear functional $\hat{f}$ on X, which is a linear extension of f and

$$\left\|\left|\hat{f}\right|\right\|_X = \|\|f\|\|_Z.$$

Proof. Let $Z = \{0\}$. Then f must be the zero function on Z. Hence, the zero function on X is a bounded linear functional on X with same norm as f that is a linear extension of f. Hence, the theorem holds.

Assume that $Z \neq \{0\}$. Define $p : X \to \mathbb{K}$ by,

$$p(x) = \|\|f\|\|_Z \|x\|.$$

Then, $|f(x)| \leq p(x)$ for all $x \in Z$. We want to show that p is a sublinear functional. Let $x, y \in X$ and α be any scalar. Then we have,

$$\begin{aligned} p(x+y) &= |||f|||_Z \,\|x+y\| \\ &\leq |||f|||_Z(\|x\| + \|y\|) \\ &= |||f|||_Z \,\|x\| + |||f|||_Z \,\|y\| \\ &= p(x) + p(y). \end{aligned}$$

And,

$$\begin{aligned} p(\alpha x) &= |||f|||_Z \,\|\alpha x\| \\ &= |\alpha| |||f|||_Z \,\|x\| \\ &= |\alpha| p(x). \end{aligned}$$

Hence, p is a sublinear functional and $|f(x)| \leq p(x)$ for all $x \in Z$. By Theorem 4.4, there exists a linear extension $\hat{f}$ of f such that $|\hat{f}(x)| \leq p(x)$ for all $x \in X$.

$$\left|\left|\left|\hat{f}\right|\right|\right|_X = \sup_{\|x\|=1} |\hat{f}(x)| \leq p(x) = |||f|||_Z \,\|x\| = |||f|||_Z.$$

Hence, $\left|\left|\left|\hat{f}\right|\right|\right|_X \leq |||f|||_Z$. Also, $\left|\left|\left|\hat{f}\right|\right|\right|_X \geq |||f|||_Z$ since $\hat{f}$ is a linear extension of f. Hence, $\left|\left|\left|\hat{f}\right|\right|\right|_X = |||f|||_Z$. Therefore, $\hat{f}$ is the desired linear extension of f. ■

Theorem 4.6 (Bounded Linear Functionals)

Let X be a non-trivial normed space and let $x_0 \neq 0$ be any element of X. Then there exists a bounded linear functional $\hat{f}$ on X such that

$$\begin{aligned} \left|\left|\left|\hat{f}\right|\right|\right| &= 1, \\ \hat{f}(x_0) &= |||x_0|||. \end{aligned}$$

Proof. Let $Z = \text{span}\,\{x_0\}$ be a subspace of X. Then any $x \in Z$ can be written as $x = \alpha x_0$ for some scalar α. Define $f : Z \to \mathbb{K}$ by,

$$f(\alpha x_0) = \alpha |||x_0||| \quad \forall \alpha \in \mathbb{K}.$$

Then, f is a linear functional on Z.

$$|f(x)| = |f(\alpha x_0)| = |\alpha| |||x_0||| = |||\alpha x_0||| = |||x||| \quad \forall x \in Z.$$

Therefore, $|||f|||_Z = 1$. By Theorem 4.5, there exists a bounded linear functional $\hat{f}$ on X such that $\left|\left|\left|\hat{f}\right|\right|\right| = 1$. Since $x_0 \in Z$, we have,

$$\hat{f}(x_0) = f(x_0) = f(1 \cdot x_0) = 1 \cdot |||x_0||| = |||x_0|||.$$

Hence, $\hat{f}$ is the desired bounded linear functional. ■

Corollary 4.7 (Norm, Zero vector)

For every x in a non-trivial normed space X,

$$\|x\| = \sup_{f \in X^*, f \neq 0} \frac{|f(x)|}{|||f|||}.$$

If for some $x_0 \in X$, $f(x_0) = 0$ for all $f \in X^*$, then $x_0 = 0$.

Proof. Let $x \in X$. If $x = 0$, then it trivially holds. Asume that $x \neq 0$. By Theorem 4.6, there exists a bounded linear functional $\hat{f}$ on X such that $\left\|\left|\hat{f}\right|\right\| = 1$ and $\hat{f}(x_0) = |||x_0|||$. Then,

$$\begin{aligned} \sup_{f \in X^*, f \neq 0} \frac{|f(x)|}{|||f|||} &\geq \frac{|\hat{f}(x)|}{\left\|\left|\hat{f}\right|\right\|} \\ &= \frac{\|x\|}{1} \\ &= 1 \end{aligned}$$

Also, $|f(x)| \leq |||f||| |||x|||$. Hence,

$$\sup_{f \in X^*, f \neq 0} \frac{|f(x)|}{|||f|||} \leq |||x|||.$$

Hence, $\sup_{f \in X^*, f \neq 0} \frac{|f(x)|}{|||f|||} = |||x|||$.

Let $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in X^*$. Then,

$$\|x_0\| = \sup_{f \in X^*, f \neq 0} \frac{|f(x_0)|}{|||f|||} = 0.$$

Therefore, $x_0 = 0$. ■

§4.3 The Adjoint Operator

Definition of the adjoint operator. Suppose we pick any bounded linear functional g on Y (that is, $g \in Y'$). We can define a linear functional f on X by

$$f(x) \;=\; g(Tx), \quad x \in X. \tag{4.7}$$

Clearly, f is linear because it is the composition of the linear map T with the linear functional g. Moreover, f is bounded since

$$|f(x)| \;=\; |g(Tx)| \;\leq\; \|g\| \, \|Tx\| \;\leq\; \|g\| \, \|T\| \, \|x\|.$$

By taking the supremum over all x with $\|x\| = 1$, we find

$$\|f\| \;\leq\; \|g\| \; |||T|||. \tag{4.8}$$

Hence f belongs to X'. Because $g \in Y'$ was arbitrary, the formula (4.7) actually defines an operator

$$T^\times : Y' \;\longrightarrow\; X',$$

called the *adjoint operator* of T. Thus we have the diagram:

$$\begin{aligned} T: \quad & X \longrightarrow Y, \\ T^\times: \quad & Y' \longrightarrow X'. \end{aligned}$$

Note that $T^\times$ acts on elements of Y', whereas T itself is defined on X.

Definition 4.7 (Adjoint operator $T^\times$)**.** If $T : X \to Y$ is a bounded linear operator between normed spaces X and Y, then the *adjoint operator* $T^\times : Y' \;\longrightarrow\; X'$ is given by

$$(T^\times g)(x) \;=\; g(Tx), \quad g \in Y',\; x \in X. \tag{4.9}$$

Our main objective is to show that $T^\times$ is itself linear and bounded, and has the same norm as T.

Theorem 4.8 (Norm of the adjoint)
For the adjoint operator $T^\times$ from Definition 4.7, we have:

$$\|T^\times\| \;=\; \|T\|.$$

Proof. First, it is straightforward to check that $T^\times$ is linear: if $g_1, g_2 \in Y'$ and α, β are scalars, then

$$(T^\times(\alpha\, g_1 + \beta\, g_2))(x) \;=\; (\alpha\, g_1 + \beta\, g_2)(Tx) \;=\; \alpha\, g_1(Tx) + \beta\, g_2(Tx) \;=\; \alpha\,(T^\times g_1)(x) \;+\; \beta\,(T^\times g_2)(x).$$

Next, from (4.8) we know that for any $g \in Y'$, if we set $f = T^\times g$, then

$$\|f\| \;\leq\; \|g\|\,\|T\|.$$

Taking the supremum over all g with $\|g\| = 1$ shows that

$$\|T^\times\| \;\leq\; \|T\|. \tag{4.10}$$

To prove the reverse inequality, $\|T^\times\| \geq \|T\|$, we use Theorem 4.3-3: given any nonzero $x_0 \in X$, there exists a functional $g_0 \in Y'$ with $\|g_0\| = 1$ such that

$$g_0(Tx_0) \;=\; \|Tx_0\|.$$

Define $f_0 = T^\times g_0$, so $f_0(x_0) = g_0(Tx_0) = \|Tx_0\|$. It follows that

$$\|Tx_0\| \;=\; |f_0(x_0)| \;\leq\; \|f_0\|\,\|x_0\| \;=\; \|T^\times g_0\|\,\|x_0\| \;\leq\; \|T^\times\|\,\|g_0\|\,\|x_0\| \;=\; \|T^\times\|\,\|x_0\|.$$

Since $\|Tx_0\| \leq \|T\|\,\|x_0\|$ always, we deduce

$$\|T\| \;=\; \sup_{x_0 \neq 0} \frac{\|Tx_0\|}{\|x_0\|} \;\leq\; \|T^\times\|.$$

Combining this with (4.10) yields $\|T^\times\| = \|T\|$, as desired. ■

§4.4 Additional Properties of the Adjoint Operator

Let $T : X \to Y$ be a bounded linear operator between normed spaces X and Y. Recall from earlier that the *adjoint operator*

$$T^\times : Y' \;\longrightarrow\; X'$$

is defined by

$$(T^\times g)(x) \;=\; g(Tx), \quad \text{for all } g \in Y',\ x \in X.$$

In finite dimensions (over $\mathbb{R}$), this corresponds to taking the *transpose* of the matrix that represents T. Over $\mathbb{C}$, one must distinguish $T^\times$ from the usual "conjugate transpose" (the Hilbert-adjoint), as explained below.

Linear Operations on the Adjoint

If $S, T \in B(X, Y)$ (the space of bounded linear maps $X \to Y$), then:

$$(S + T)^\times = S^\times + T^\times, \tag{9}$$
$$(\alpha T)^\times = \alpha\, T^\times, \tag{10}$$

for any scalar α. These properties mirror the usual linearity of the adjoint.

Product of Operators

Suppose $T \in B(X,Y)$ and $S \in B(Y,Z)$ for normed spaces X,Y,Z. Then their product $ST \in B(X,Z)$ has adjoint satisfying

$$(ST)^\times \;=\; T^\times S^\times, \tag{11}$$

as shown below:

$$X \xrightarrow{T} Y \xrightarrow{S} Z \quad \longmapsto \quad X' \xleftarrow{T^\times} Y' \xleftarrow{S^\times} Z'.$$

Inverse of the Adjoint

If T is invertible, i.e. $T \in B(X,Y)$ and $T^{-1} \in B(Y,X)$, then the adjoint $T^\times$ is invertible too, and

$$(T^\times)^{-1} \;=\; (T^{-1})^\times, \tag{12}$$

meaning invertibility of T carries over to its adjoint on dual spaces.

Adjoint vs. Hilbert-Adjoint

In the special case where $X = H_1$ and $Y = H_2$ are Hilbert spaces, the operator T^* defined in Section 3.9 is the *Hilbert-adjoint* of T. Meanwhile, $T^\times$ is again the map $Y' \to X'$. Via the Riesz representation theorem, each bounded functional on a Hilbert space corresponds to a unique vector, and one can show that

$$T^* \;=\; A_1 \, T^\times \, A_2^{-1},$$

where A_1 and A_2 are conjugate-linear isometries implementing the Riesz representation on H_1' and H_2' respectively. Thus for all $x \in H_1$ and $y_0 \in H_2$,

$$\langle Tx,\, y_0\rangle_{H_2} \;=\; \langle x,\, T^* y_0\rangle_{H_1}.$$

Since T^* is essentially $T^\times$ composed with conjugate-linear isomorphisms, in a complex Hilbert space we get the usual formula $(\alpha T)^* = \overline{\alpha}\, T^*$, whereas in the adjoint on normed spaces we have $(\alpha T)^\times = \alpha\, T^\times$. Moreover, $\|T^*\| = \|T\|$ follows immediately (Theorem 3.9-2) by combining $\|T^\times\| = \|T\|$ with the isometric nature of the Riesz maps A_1 and A_2.

Remarks

- $T^\times$ acts on the dual space Y' (the space of functionals on Y) rather than on Y itself.
- T^* in a Hilbert space is the "usual" adjoint that satisfies $\langle Tx,\, y\rangle = \langle x,\, T^* y\rangle$.
- In finite dimensions over $\mathbb{R}$, $T^\times$ is represented by the transpose of the matrix of T, whereas in the complex case, T^* is given by the conjugate transpose (also called the Hermitian adjoint).

§4.5 Reflexive Spaces

We previously examined the notion of *algebraic reflexivity* in Section 2.8, where we introduced the *second algebraic dual space* $X^{**} = (X^*)^*$ of a vector space X. There, X was called *algebraically reflexive* if the *canonical map*

$$C : X \;\longrightarrow\; X^{**}$$

is surjective. This map is defined by sending each $x \in X$ to the functional $g_x \in X^{**}$ given by

$$g_x(f) = f(x), \quad (f \in X^*).$$

Hence for each $x \in X$, g_x is the linear functional that evaluates any $f \in X^*$ at x. In the finite-dimensional case, X is indeed algebraically reflexive (Theorem 2.27).

Here, we focus on *normed* spaces. Let X be a normed space with dual X', as defined in Section 2.10–3, and let $(X')'$ (the dual of X') be denoted by X''. We call X'' the *second dual* or *bidual* of X. For each fixed $x \in X$, define g_x on X' by

$$g_x(f) = f(x), \quad (f \in X').$$

Although this resembles the algebraic version (1), here f is necessarily a bounded functional. That g_x is itself a bounded linear functional follows from the next lemma.

Lemma 4.9 (Norm of g_x)

If x is any element of a normed space X, the map g_x defined by (2) is in X'' and satisfies

$$\|g_x\| \;=\; \|x\|.$$

Proof. Linearity of g_x can be verified by the same argument used in Section 2.8. Observe that

$$\|g_x\| \;=\; \sup_{f \in X',\, f \neq 0} \frac{|g_x(f)|}{\|f\|} \;=\; \sup_{f \in X',\, f \neq 0} \frac{|f(x)|}{\|f\|} \;=\; \|x\|,$$

where the last equality follows from a standard dual-norm argument (Corollary 4.7). ■

From each $x \in X$, we thus get a unique $g_x \in X''$, and we can define the *canonical map*

$$C : X \longrightarrow X'', \quad x \mapsto g_x.$$

Since it is injective and preserves norms (as seen below), C is an *isometric embedding*.

Lemma 4.10 (Canonical mapping)

The map C from (5) is an isomorphism of X onto its image $\mathcal{R}(C) \subset X''$, meaning it is linear, norm-preserving (isometric), and bijective onto $\mathcal{R}(C)$.

Proof. Linearity follows by checking

$$g_{\alpha x+\beta y}(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha\, g_x(f) + \beta\, g_y(f),$$

implying $g_{\alpha x+\beta y} = \alpha\, g_x + \beta\, g_y$. Furthermore,

$$\|g_x - g_y\| = \|g_{x-y}\| = \|\, x - y\|,$$

so C is an isometry, which forces injectivity. Surjectivity onto $\mathcal{R}(C)$ is clear by definition of "range," so C is indeed an isomorphism onto its image. ■

In other words, X can be *embedded* into X''. However, C need not be surjective onto all of X''. We call X *reflexive* precisely when C *is* onto X'':

Definition 4.8 (Reflexivity). A normed space X is called *reflexive* if its canonical map

$$C : X \;\to\; X'', \quad x \;\mapsto\; g_x$$

is onto, i.e. $\mathcal{R}(C) = X''$.

This concept dates back to H. Hahn (1927) and was named "reflexivity" by E. R. Lorch (1939). If X is reflexive, then by Lemma 4.10, X and X'' are isomorphic and isometric. However, the converse does not hold in full generality, as shown by R. C. James.

Moreover, reflexivity does not simply follow from completeness, although there is a partial converse:

Theorem 4.11 (Completeness implies Banach)
If a normed space X is reflexive, then X must be complete (hence is a Banach space).

Proof. Since X'' is the dual of X', it is known (Theorem 2.10-4) to be complete. If X is reflexive, we have $X'' = \mathcal{R}(C)$, and the isomorphism then yields completeness of X from that of X''. ■

Examples and Non-Examples. Every finite-dimensional normed space is reflexive (Theorem 2.27), since all linear functionals are bounded in finite dimensions, and the algebraic reflexivity implies the normed version. In particular, $\mathbb{R}^n$ is reflexive.

For $1 < p < \infty$, the spaces ℓ^p and $L^p[a, b]$ can also be shown to be reflexive. By contrast, many classical Banach spaces fail to be reflexive, e.g. $C[a, b]$, $L^1[a, b]$, and ℓ^∞. Nonetheless, every Hilbert space is reflexive:

Theorem 4.12 (Hilbert spaces are reflexive)
Every Hilbert space H is reflexive.

Sketch of proof. We show the canonical map $C : H \to H''$ is onto. For any $g \in H''$, one wants an $x \in H$ so that $g = Cx$. A key step uses the Riesz representation theorem to construct a conjugate-linear isometric map $A : H' \to H$ and then carry out a second representation argument to identify g with evaluation at some $x \in H$. Thus each g is in the range of C. ■

Remark 4.2. A normed space X that is separable but has a nonseparable dual X' cannot be reflexive. Indeed, if X were reflexive, then $X'' \cong X$ by Lemma 4.10, and thus the separability of X would force X'' (and hence X') to be separable as well, which is a contradiction.

Example 4.1. ℓ^1 is not reflexive.

Proof. ℓ^1 is separable, but its dual ℓ^∞ is not separable. Hence ℓ^1 fails to be reflexive. ■

We next prove a lemma (often illustrated by a geometric argument in $\mathbb{R}^3$) that enables us to construct a functional with norm 1 that vanishes on a given closed subspace.

Lemma 4.13 (Existence of a functional)
Suppose Y is a proper closed subspace of a normed space X, and let $x_0 \in X \setminus Y$. Define

$$\delta = \inf_{y \in Y} \|y - x_0\|.$$

Then there is an $f \in X'$ with $\|f\| = 1$, such that $f(y) = 0$ for all $y \in Y$ and $f(x_0) = \delta$.

Sketch of proof. Let Z be the linear span of Y and x_0. One first defines a linear functional f on Z by

$$f(y + \alpha\, x_0) = \alpha\, \delta,$$

and checks f is bounded with norm at most 1, and that f attains the value δ at x_0 while vanishing on Y. A standard argument shows $\|f\| \geq 1$, so in fact $\|f\| = 1$ on Z. Finally, the Hann-Banach Theorem for Normed Spaces extends f from Z to all of X without increasing its norm, giving the required functional. ■

With Lemma 4.13 in place, we can establish a well-known separability criterion:

Theorem 4.14 (Separability)
If the dual X' of a normed space X is separable, then X itself must also be separable.

Proof. Assume X' is separable. Consider its unit sphere

$$U' = \{ f \in X' : \|f\| = 1\}.$$

This set inherits a countable dense subset, say $\{f_n\}_{n=1}^{\infty}$, because X' is separable. Each f_n has norm 1, so by definition of the supremum that yields $\|f_n\| = 1$, we can choose $x_n \in X$ with $\|x_n\| = 1$ and $|f_n(x_n)| \geq \frac{1}{2}$.

Let Y be the closure of the linear span of $\{x_n\}$. This subspace Y is separable because it has a countable dense set (rational combinations of the x_n). If Y were strictly smaller than X, then by Lemma 4.13 there would be some $f \in X'$ with $\|f\| = 1$ that is zero on Y but nonzero at a point outside Y. In particular, $f(x_n) = 0$ for every n, while $|f_n(x_n)| \geq \frac{1}{2}$. Hence

$$\frac{1}{2} \leq |f_n(x_n)| = |f_n(x_n) - f(x_n)| \leq \|f_n - f\| \, \|x_n\| = \|f_n - f\|.$$

So $\|f_n - f\| \geq \frac{1}{2}$ for all n. But this contradicts the assumption that $\{f_n\}$ is dense in the unit sphere (since f is also on that sphere with $\|f\| = 1$). Therefore, our assumption $Y \neq X$ must fail, so $Y = X$.

Hence X is generated (densely) by the countable set $\{x_n\}$, proving that X is separable. ■

§4.6 Baire's Category Theorem in Complete Metric Spaces

Definition 4.9 (Category in a Metric Space)**.** Let X be a metric space. A subset $M \subset X$ is called

- *nowhere dense* (or *rare*) in X if the closure of M, denoted $\overline{M}$, has empty interior (see also Sec. 1.3);
- *meager* (or *of first category*) in X if it can be expressed as a countable union of nowhere dense sets;
- *nonmeager* (or *of second category*) in X if it is not meager in X.

Theorem 4.15 (Baire's Category Theorem for Complete Metric Spaces)

Suppose X is a nonempty, complete metric space. Then X cannot be meager in itself (i.e. X is of second category in X).

Consequently, if $X \neq \varnothing$ is complete and can be written as

$$X = \bigcup_{k=1}^{\infty} A_k \quad \text{(each } A_k \text{ closed in } X\text{)},$$

then at least one of the closed sets A_k contains a nonempty open subset of X.

Proof. **Contradiction argument.** Assume instead that X is meager in itself, despite being nonempty and complete. Then there exist sets M_k, each nowhere dense in X, such that

$$X = \bigcup_{k=1}^{\infty} M_k.$$

We will construct a Cauchy sequence (p_k) within X converging to some $p \in X$ that does not lie in any M_k, contradicting $X = \bigcup M_k$.

Step 1: Choosing a point and ball outside each M_k. Since M_1 is nowhere dense, its closure $\overline{M_1}$ has no interior, so there is an open set in X that stays entirely outside $\overline{M_1}$. Call that open set M_1^c, and pick $p_1 \in M_1^c$ along with a small ball

$$B_1 = B(p_1; \varepsilon_1) \subset M_1^c.$$

For M_2, again its closure $\overline{M_2}$ lacks any nonempty open set, so it cannot include $B(p_1; \frac{1}{2}\varepsilon_1)$ in its entirety. Hence there is a point p_2 and a ball

$$B_2 = B(p_2; \varepsilon_2) \subset M_2^c \cap B(p_1; \tfrac{1}{2}\varepsilon_1),$$

with $\varepsilon_2 < \frac{1}{2}\varepsilon_1$.

Step 2: Inductive construction. Continuing inductively, we obtain a sequence of balls

$$B_k = B(p_k; \varepsilon_k)$$

where each B_k stays outside M_k and also lies inside the previous ball (shrunk by half),

$$B_k \subset M_k^c \quad \text{and} \quad B_k \subset B(p_{k-1}; \tfrac{1}{2}\varepsilon_{k-1}).$$

Additionally, choose ε_k so that $\varepsilon_k < 2^{-k}$.

Step 3: Cauchy sequence and limit point. Because ε_k tends to zero (faster than 2^{-k}), the centers p_k form a Cauchy sequence in the complete space X. Let $p_k \to p \in X$. By construction, p remains within each ball B_k, and thus p never belongs to M_k since $B_k \subset M_k^c$. This contradicts the assumption that every point of X lies in some M_k.

Thus X cannot be meager in itself. The statement about closed sets A_k follows by a similar argument, ensuring that one A_k must admit a nonempty open subset. ■

§4.7 Strong and Weak Convergence

§4.7.i Strong Convergence

Definition 4.10 (Strong convergence). A sequence (x_n) in a normed space X *converges strongly* (or *in the norm*) to an element $x \in X$ if

$$\lim_{n\to\infty} \| x_n - x \| = 0.$$

We typically write $x_n \to x$ or $\lim x_n = x$, and we refer to x as the *strong limit* of (x_n).

Intuitively, strong convergence means that x_n approaches x in the norm sense, so $\|x_n - x\|$ becomes arbitrarily small.

§4.7.ii Weak Convergence

Definition 4.11 (Weak convergence). A sequence (x_n) in a normed space X *converges weakly* to an element $x \in X$ if, for every bounded linear functional f in the dual space X', we have

$$\lim_{n\to\infty} f(x_n) \;=\; f(x).$$

We denote this by $x_n \xrightarrow{w} x$, and call x the *weak limit* of (x_n).

Thus, weak convergence can be understood as pointwise convergence of the sequence $(f(x_n))$ for all f in X'.

Lemma 4.16 (Basic properties of weak convergence)
Suppose (x_n) is a weakly convergent sequence in a normed space X, so $x_n \xrightarrow{w} x$. Then:

(a) The weak limit x is unique.

(b) Every subsequence of (x_n) also converges weakly to x.

(c) The sequence $\|x_n\|$ is bounded.

Proof. **(a)** If we also had $x_n \xrightarrow{w} y$ for some $y \in X$, then for all $f \in X'$ the sequence $f(x_n)$ would converge to both $f(x)$ and $f(y)$. Uniqueness of the limit for scalars implies $f(x) = f(y)$. Since this holds for all $f \in X'$, one deduces $x = y$ by a standard separation argument (Corollary 4.3-4).
(b) Since $f(x_n)$ converges (as a sequence of scalars) for every $f \in X'$, the same limit must apply to every subsequence, hence each subsequence converges weakly to x.
(c) Each $(f(x_n))$ is a convergent (thus bounded) sequence of scalars. By the Uniform Boundedness Theorem, the family of functionals $g_n \in X''$ defined by $g_n(f) := f(x_n)$ has bounded norms $\|g_n\|$. But from the canonical identification (Section 4.6-1), $\|g_n\| = \|x_n\|$. Hence $(\|x_n\|)$ is bounded. ∎

§4.7.iii Comparison of Strong and Weak Convergence

Theorem 4.17 (Strong vs. weak convergence)
Let (x_n) be a sequence in a normed space X. Then the following statements hold:

(a) If $x_n \to x$ strongly (i.e. $\|x_n - x\| \to 0$), then $x_n \xrightarrow{w} x$ weakly, with the same limit.

(b) The converse does not hold in general: weak convergence need not imply strong convergence.

(c) If $\dim X < \infty$, then weak convergence *does* imply strong convergence.

Proof. **(a)** Strong convergence $\|x_n - x\| \to 0$ directly forces $f(x_n) \to f(x)$ for every $f \in X'$, simply because

$$|f(x_n) - f(x)| \;=\; |f(x_n - x)| \;\le\; \|f\|\,\|x_n - x\| \;\longrightarrow\; 0.$$

Hence $x_n \xrightarrow{w} x$.
(b) A classical example arises in a Hilbert space with an orthonormal sequence (e_n). One can show $e_n \xrightarrow{w} 0$ (since $\langle e_n, z\rangle \to 0$ for every $z \in H$), but $\|e_n - 0\| = 1$ never goes to 0. So weak convergence need not imply strong convergence in infinite dimensions.
(c) In a finite-dimensional space X, all norms are equivalent, and a finite number of linear functionals (the "dual basis") can distinguish any point in X. If x_n converges weakly to x, then componentwise (with respect to a basis) we see $\|x_n - x\| \to 0$, hence strong convergence as well. ∎

Moreover, there are particular infinite-dimensional spaces—such as ℓ^1 under a result of I. Schur (1921)—in which strong and weak convergence coincide. However, this is the exception rather than the rule.

Example 4.2 (Hilbert space)**.** If H is a Hilbert space, then $x_n \xrightarrow{w} x$ if and only if

$$\langle x_n,\, z\rangle \;\longrightarrow\; \langle x, z\rangle \quad \text{for all } z \in H.$$

This equivalence relies on the Riesz Representation Theorem, which identifies every functional $f \in H'$ with an inner product $\langle \cdot,\, z\rangle$ for some unique $z \in H$.

Example 4.3 ($\boldsymbol{\ell^p}$ **with** $\mathbf{1 < p < \infty}$)**.** For a sequence (x_n) in ℓ^p, write $x_n = (\xi_j^{(n)})_{j=1}^{\infty}$ and $x = (\xi_j)_{j=1}^{\infty}$. Then x_n converges weakly to x precisely if

(A) $\|x_n\|$ is bounded, and

(B) For every coordinate j, $\xi_j^{(n)} \to \xi_j$.

Since the dual of ℓ^p is ℓ^q (where $1/p + 1/q = 1$), the bounded linear functionals on ℓ^p come from sequences in ℓ^q whose standard basis picks out individual coordinates.

Lemma 4.18 (Weak convergence criterion)
In a normed space X, a sequence (x_n) converges weakly to x if and only if

(A) $(\|x_n\|)$ remains bounded, and

(B) $f(x_n) \to f(x)$ for every f in some *total* subset $M \subset X'$, where "total" means that the linear span of M is dense in X'.

Proof. (⇒) From Lemma 4.16(c) we get that $\|x_n\|$ is bounded. And trivially, if $f(x_n) \to f(x)$ for *all* $f \in X'$, it is true in particular for a total subset $M \subset X'$.
(⇐) Assume (A) and (B). Fix any $f \in X'$. By denseness, there is a sequence $(f_j) \subset \operatorname{span} M$ approaching f in the norm of X'. Moreover, from (B) one sees $f_j(x_n) \to f_j(x)$. Carefully choosing large j and n makes $|f(x_n) - f(x)|$ as small as desired. Thus $f(x_n) \to f(x)$ for all $f \in X'$, i.e. $x_n \xrightarrow{w} x$. ■

§4.8 Convergence of Sequences of Operators and Functionals

- **Uniform (operator) convergence**: $\|T_n - T\| \to 0$ in the operator norm on $B(X,Y)$.
- **Strong (operator) convergence**: for each fixed $x \in X$, $\|T_n x - Tx\| \to 0$ in Y.
- **Weak (operator) convergence**: for each $x \in X$, $\{T_n x\}$ converges *weakly* in Y; that is,

$$f(T_n x) \longrightarrow f(Tx) \quad \text{for all } f \in Y'.$$

Definition 4.12 (Operator convergence)**.** Let X and Y be normed spaces, and let $\{T_n\} \subset B(X,Y)$ be a sequence of bounded operators. We say (T_n) *converges uniformly* (or in norm) to an operator $T \in B(X,Y)$ if

$$\| T_n - T \| \longrightarrow 0 \quad \text{in } B(X,Y).$$

It *converges strongly* to T if

$$\| T_n x - Tx \| \longrightarrow 0 \quad \text{for every } x \in X.$$

It *converges weakly* to T if

$$f(T_n x) \longrightarrow f(Tx) \quad \text{for all } x \in X \text{ and } f \in Y'.$$

One shows readily that

$$(\text{uniform convergence}) \implies (\text{strong convergence}) \implies (\text{weak convergence}),$$

but the converses fail in general, as illustrated by classical examples in ℓ^2:

Example 4.4 (Strong, but not uniform, operator convergence)**.** Define $T_n : \ell^2 \to \ell^2$ by

$$T_n(x_1, x_2, \dots) = (0, 0, \dots, 0, x_{n+1}, x_{n+2}, \dots),$$

where the first n entries are replaced by zeros. Each T_n is bounded with operator norm $\|T_n\| = 1$. For any fixed $x \in \ell^2$, the tail $(x_{n+1}, x_{n+2}, \dots)$ goes to zero in ℓ^2, so $T_n x \to 0$. Hence $\{T_n\}$ converges *strongly* to the zero operator. It does *not* converge uniformly, however, since $\|T_n - 0\| = \|T_n\| = 1 \not\to 0$.

Example 4.5 (Weak, but not strong, operator convergence)**.** Again in ℓ^2, define $T_n : \ell^2 \to \ell^2$ by shifting the first n components to zero:

$$T_n(x_1, x_2, \dots) = (0, 0, \dots, 0, x_1, x_2, \dots).$$

Then for any linear functional $f \in (\ell^2)' \cong \ell^2$ (by Riesz representation), we find $f(T_n x) \to 0$ by a tail argument in the inner product, implying $T_n x$ converges *weakly* to 0 in ℓ^2. However, choosing $x = (1, 0, 0, \dots)$ shows that $\|T_n x\|$ does not tend to 0, so $\{T_n\}$ cannot converge strongly to 0.

Definition 4.13 (Strong and weak* convergence for functionals)**.** Suppose $\{f_n\}$ is a sequence of bounded linear functionals on a normed space X.

- **Strong convergence:** There is an $f \in X'$ such that $\|f_n - f\| \to 0$. One writes $f_n \to f$ in norm.
- **Weak* convergence:** There is an $f \in X'$ such that $f_n(x) \to f(x)$ for all $x \in X$. We denote this $f_n \xrightarrow{w^*} f$.

These two notions coincide exactly with the strong vs. weak operator convergence in the finite-dimensional codomain $\mathbb{F}$.

Lemma 4.19 (Strong limit is bounded)

Suppose X is a Banach space, Y is a normed space, and $T_n \in B(X,Y)$ are such that $T_n x \to Tx$ pointwise in Y. Then T is automatically a bounded linear operator in $B(X,Y)$.

Proof. Linearity of T follows from that of T_n. Boundedness is ensured by the Uniform Boundedness Theorem: since X is complete, the maps T_n cannot all be large on different x's without violating uniform boundedness. Formally, $\sup_n \|T_n\| \le c < \infty$ implies $\|Tx\| \le c\|x\|$ for each x, so $T \in B(X,Y)$. ∎

Theorem 4.20 (Strong operator convergence criterion)

Let X and Y be Banach spaces, and let $\{T_n\} \subset B(X,Y)$. Then $T_n \to T$ strongly (i.e. $T_n x \to Tx$ for all $x \in X$) *if and only if*:

(A) $\{\|T_n\|\}$ remains bounded;

(B) For every x in a *total* subset $M \subset X$, the sequence $\{T_n x\}$ is Cauchy in Y.

Proof. *(⇒)* If $T_n x \to Tx$ for all x, by uniform boundedness we get a bound on $\|T_n\|$, giving (A). And trivially $(T_n x)$ is Cauchy in Y for every x (particularly those x in M).

(⇐) Assume (A) and (B). Fix $x \in X$, and choose y in the span of M with $\|x-y\| < \varepsilon/(3c)$. Since $(T_n y)$ is Cauchy in Y, there is N so that for $m,n > N$, $\|T_n y - T_m y\| < \varepsilon/3$. A short argument bounding $\|T_n x - T_m x\|$ by $\|T_n x - T_n y\| + \|T_n y - T_m y\| + \|T_m y - T_m x\|$ shows $(T_n x)$ is also Cauchy. Completeness of Y then ensures $T_n x \to Tx$, and Lemma 4.19 confirms $T \in B(X,Y)$. ∎

Corollary 4.21 (Weak* convergence of functionals)

For a sequence of bounded functionals $\{f_n\}$ on a Banach space X, the condition $f_n \xrightarrow{w^*} f$ holds if and only if

(A) $\{\|f_n\|\}$ is bounded,

(B) For every x in some total subset $M \subset X$, $\{f_n(x)\}$ is Cauchy in the scalar field.

In other words, this "weak*" convergence of functionals parallels the strong convergence of operators in a one-dimensional codomain.

§4.9 Open Mapping Thoerem

If V is a vector space over the field $\mathbb{F}$ and $A \subseteq V$, then for $v \in V$ and $\alpha \in \mathbb{F}$, the sets $A + v$ and αA are defined as

$$A + v = \{a + v \mid a \in A\} \text{ and } \alpha A = \{\alpha a \mid a \in A\}. \tag{4.11}$$

Lemma 4.22

A bounded linear operator T from a Banach space X onto another Banach space Y has the property that $T(B(0,1))$ contains an open ball centred at $0 \in Y$.

Proof. Let $B_n = B(0, 2^{-n})$. So $B_1 = B\left(0, \frac{1}{2}\right)$. For any $x \in X$, $x \in kB_1$ for some $k > 2\|x\|$. Therefore, $X = \bigcup_{k\in\mathbb{N}} kB_1$. Since T is surjective and linear,

$$Y = T(X) = T\left(\bigcup_{k\in\mathbb{N}} kB_1\right) = \bigcup_{k=1}^{\infty} kT(B_1). \tag{4.12}$$

$kT(B_1) \subseteq \overline{kT(B_1)} \subseteq Y$, so

$$Y = \bigcup_{k=1}^{\infty} kT(B_1) = \bigcup_{k=1}^{\infty} \overline{kT(B_1)}. \tag{4.13}$$

Since Y is complete, by Baire's Category Theorem for Complete Metric Spaces, $\overline{kT(B_1)}$ contains an open ball $B(a_0, r)$ for some k. Let $y_0 = \frac{a_0}{k}$ and $\varepsilon = \frac{r}{k}$.

$$B(a_0, r) \subseteq \overline{kT(B_1)} = k\overline{T(B_1)} \implies B(y_0, \varepsilon) = B\left(\frac{a_0}{k}, \frac{r}{k}\right) \subseteq \overline{T(B_1)}. \tag{4.14}$$

Let $B^* = B(y_0, \varepsilon)$. By translating, $B^* - y_0 = B(0, \varepsilon) \subseteq \overline{T(B_1)} - y_0$. Now we want to show that $B^* - y_0 \subseteq \overline{T(B_0)}$. It suffices to show that $\overline{T(B_1)} - y_0 \subseteq \overline{T(B_0)}$.

Let $y \in \overline{T(B_1)} - y_0$. Then $y + y_0 \in \overline{T(B_1)}$. y_0 is also in $\overline{T(B_1)}$ as $B(y_0, \varepsilon) \subseteq \overline{T(B_1)}$. Then there are sequences $u_n, v_n \in T(B_1)$ converging to $y + y_0$ and y_0 respectively. Since $u_n, v_n \in T(B_1)$, $u_n = Tw_n$ and $v_n = Tz_n$ for some $w_n, z_n \in B_1$.

$$\|w_n - z_n\| \leq \|w_n\| + \|z_n\| < \frac{1}{2} + \frac{1}{2} = 1 \implies w_n - z_n \in B_0. \tag{4.15}$$

$T(w_n - z_n) = u_n - v_n$ converges to $y + y_0 - y_0 = y$. Therefore, $y \in \overline{T(B_0)}$, proving that $\overline{T(B_1)} - y_0 \subseteq \overline{T(B_0)}$. As a result, $B^* - y_0 = B(0, \varepsilon) \subseteq \overline{T(B_0)}$.

$B_n = B(0, 2^{-n}) = 2^{-n}B(0,1) = 2^{-n}B_0$, so

$$\overline{T(B_n)} = \overline{T(2^{-n}B_0)} = 2^{-n}\overline{T(B_0)}. \tag{4.16}$$

Since $B(0, \varepsilon) \subseteq \overline{T(B_0)}$, scaling by 2^{-n} gives us $B(0, 2^{-n}\varepsilon) \subseteq \overline{T(B_n)}$. Now, finally, we shall prove that $B\left(0, \frac{1}{2}\varepsilon\right) \subseteq T(B_0)$. Take $y \in B\left(0, \frac{1}{2}\varepsilon\right)$. $B(0, 2^{-1}\varepsilon) \subseteq \overline{T(B_1)}$, so $y \in \overline{T(B_1)}$. Then there exists $v = Tx_1 \in T(B_1)$ such that

$$\|y - v\| = \|y - Tx_1\| < \frac{\varepsilon}{4}. \tag{4.17}$$

So $y - Tx_1 \in B(0, 2^{-2}\varepsilon) \subseteq \overline{T(B_2)}$. Similar as before, there exists $x_2 \in B_2$ such that

$$\|y - Tx_1 - Tx_2\| < \frac{\varepsilon}{8}. \tag{4.18}$$

Continuing this way, in the n-th step, we shall find $x_n \in B_n$ such that

$$\left\|y - \sum_{i=1}^{n} Tx_i\right\| < \frac{\varepsilon}{2^{n+1}}. \tag{4.19}$$

$x_i \in B_i$, so $\|x_i\| < 2^{-i}$. Set $z_n = x_1 + x_2 + \cdots + x_n$. We claim that (z_n) is a Cauchy sequence. Fix $\varepsilon > 0$ and take $N \in \mathbb{N}$ such that $2^N\varepsilon > 1$. Now, for $n > m \geq N$,

$$\|z_n - z_m\| = \left\|\sum_{i=m+1}^{n} x_i\right\| \leq \sum_{i=m+1}^{n} \|x_i\| < \sum_{i=m+1}^{n} 2^{-i} = 2^{-m} \leq 2^{-N} < \varepsilon. \tag{4.20}$$

So (z_n) is a Cauchy sequence, and hence it converges to some $x \in X$ since X is complete.

$$\|x\| = \left\|\lim_{n\to\infty}\sum_{i=1}^{n} x_i\right\| = \lim_{n\to\infty}\left\|\sum_{i=1}^{n} x_i\right\| \le \lim_{n\to\infty}\sum_{i=1}^{n}\|x_i\| = \sum_{i=1}^{\infty}\|x_i\|. \tag{4.21}$$

Therefore,

$$\|x\| \le \|x_1\| + \sum_{i=2}^{\infty}\|x_i\| < \frac{1}{2} + \sum_{i=2}^{\infty}\|x_i\| \le \frac{1}{2} + \sum_{i=2}^{\infty} 2^{-i} = 1 \implies x \in B_0 = B(0,1). \tag{4.22}$$

Since T is bounded, it is continuous. z_n converges to x, so Tz_n must converge to Tx due to continuity. However, from (4.19),

$$\|y - Tz_n\| = \left\|y - \sum_{i=1}^{n} Tx_i\right\| < \frac{\varepsilon}{2^{n+1}} \tag{4.23}$$

gives us that Tz_n converges to y. Therefore, $Tx = y$. Since $x \in B_0$, $y \in T(B_0)$, proving that $B\left(0, \frac{1}{2}\varepsilon\right) \subseteq T(B_0) = T(B(0,1))$. ■

Theorem 4.23 (Open Mapping Theorem)

A bounded linear map T from a Banach space X onto another Banach space Y is an open mapping, i.e. it maps open sets to open sets.

Proof. Let $A \subseteq X$ be open. We need to show that $T(A) \subseteq Y$ is open. Take $y = Tx \in T(A)$. We need to find $r > 0$ such that $B(y, r) \subseteq T(A)$.

A is open and $x \in A$, so $B(x, \varepsilon) \subseteq A$ for some $\varepsilon > 0$.

$$B(x,\varepsilon) \subseteq A \implies B(0,\varepsilon) \subseteq A - x \implies B(0,1) \subseteq \frac{1}{\varepsilon}(A - x). \tag{4.24}$$

By Lemma 4.22, $T(B(0,1))$ contains an open ball $B(0, r_0)$. $T(B(0,1))$ is contained in $T\left(\frac{1}{\varepsilon}(A - x)\right)$. Therefore,

$$B(0, r_0) \subseteq T\left(\frac{1}{\varepsilon}(A - x)\right) = \frac{1}{\varepsilon}T(A - x) = \frac{1}{\varepsilon}[T(A) - Tx] = \frac{1}{\varepsilon}[T(A) - y]. \tag{4.25}$$

Therefore,

$$\varepsilon B(0, r_0) \subseteq T(A) - y \implies \varepsilon B(0, r_0) + y = B(y, r_0\varepsilon) \subseteq T(A). \tag{4.26}$$

Since $y \in T(A)$ was arbitrary, $T(A)$ is open. ■

Corollary 4.24 (Bounded Inverse Theorem)

If a bounded linear operator $T : X \to Y$ between Banach spaces is bijective, then T^{-1} is continuous bounded.

Proof. Firstly, we will show that T^{-1} is linear. If $Tx_1 = y_1$ and $Tx_2 = y_2$, then $T(\alpha x_1 + \beta x_2) = \alpha y_1 + \beta y_2$ by the linearity of T. As a result,

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1} y_1 + \beta T^{-1} y_2. \tag{4.27}$$

Now we will show that $T^{-1} : Y \to X$ is continuous. If $A \subseteq X$ is open, $\left(T^{-1}\right)^{-1}(A) = T(A)$ is also open by Open Mapping Theorem. Therefore, T^{-1} is continuous, and hence it is bounded by Theorem 2.20. ■

§4.10 Closed Linear Operators

Lemma 4.25

Let X and Y be Banach spaces. Then $X \times Y$ is also a Banach space with respect to the norm

$$\|(x, y)\| = \|x\| + \|y\| . \tag{4.28}$$

Proof. One can check that the abovementioned norm on $X \times Y$ satisfies the properties of a norm. Let $z_n = (x_n, y_n)$ be a Cauchy sequence in $X \times Y$. Given $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that for $m, n \geq N$,

$$\|z_n - z_m\| = \|x_n - x_m\| + \|y_n - y_m\| < \varepsilon \implies \|x_n - x_m\| < \varepsilon \text{ and } \|y_n - y_m\| < \varepsilon. \tag{4.29}$$

So, both x_n and y_n are Cauchy sequences in X and Y, respectively. Since X and Y be Banach spaces, x_n and y_n converge to $x \in X$ and $y \in Y$ respectively. We claim that z_n converges to $z = (x, y)$.

Given $\varepsilon > 0$, there are positive integers N_1 and N_2 such that for $n \geq N_1$ and $m \geq N_2$,

$$\|x_n - x\| < \frac{\varepsilon}{2} \text{ and } \|y_m - y\| < \frac{\varepsilon}{2}. \tag{4.30}$$

Now, for $n \geq \max\{N_1, N_2\}$,

$$\|z_n - z\| = \|x_n - x\| + \|y_n - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{4.31}$$

Therefore, z_n converges to z proving that $X \times Y$ is complete. ■

Definition 4.14 (Closed Linear Operator)**.** Let X and Y be normed spaces, and $T : \mathcal{D}(T) \subseteq X \to Y$ a linear operator. T is called a **closed linear operator** if its graph

$$\mathcal{G}(T) = \{(x, Tx) \mid x \in \mathcal{D}(T)\} \tag{4.32}$$

is closed in the normed space $X \times Y$ equipped with the norm defined in (4.28).

Theorem 4.26 (Closed Graph Theorem)

Let X and Y be Banach spaces, and $T : \mathcal{D}(T) \subseteq X \to Y$ a closed linear operator. If $\mathcal{D}(T)$ is closed in X, T is bounded.

Proof. By Lemma 4.25, $X \times Y$ is a Banach space. Closed subsets of complete metric spaces are also complete. Therefore, both $\mathcal{G}(T)$ and $\mathcal{D}(T)$ are complete. Let $P : \mathcal{G}(T) \to \mathcal{D}(T)$ be the projection map.

$$P(x, Tx) = x. \tag{4.33}$$

P is linear since it maps $\alpha(x, Tx) + \beta(y, Ty)$ to $\alpha x + \beta y$. P is also bounded, since

$$\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\| . \tag{4.34}$$

P is bijective. Therefore, by Bounded Inverse Theorem, P^{-1} is bounded. So, we have $\|(x, Tx)\| \leq \|P^{-1}\| \|x\|$. Therefore,

$$\|Tx\| \leq \|Tx\| + \|x\| = \|(x, Tx)\| \leq \left\|P^{-1}\right\| \|x\| . \tag{4.35}$$

Hence, T is bounded. ■

Theorem 4.27 (Closed linear operator)

A linear operator $T : \mathcal{D}(T) \subset X \to Y$ is *closed* precisely when the condition below is satisfied:

$$\text{If } x_n \in \mathcal{D}(T) \text{ and } x_n \to x, \text{ while } Tx_n \to y, \text{ then } x \in \mathcal{D}(T) \text{ and } Tx = y.$$

Proof. ($\Rightarrow$) Suppose T is closed. By definition, its graph

$$\mathcal{G}(T) = \{(x, Tx) : x \in \mathcal{D}(T)\}$$

is a closed subset of $X \times Y$. Now assume $x_n \in \mathcal{D}(T)$ with $x_n \to x$ in X and $Tx_n \to y$ in Y. Then $(x_n, Tx_n) \to (x, y)$ in $X \times Y$. Since $\mathcal{G}(T)$ is closed and each $(x_n, Tx_n) \in \mathcal{G}(T)$, it follows that $(x, y) \in \mathcal{G}(T)$, so $x \in \mathcal{D}(T)$ and $Tx = y$.

($\Leftarrow$) Conversely, suppose that whenever $x_n \in \mathcal{D}(T)$, $x_n \to x$, and $Tx_n \to y$, we conclude $x \in \mathcal{D}(T)$ and $Tx = y$. We claim that $\mathcal{G}(T)$ is closed. Take any sequence $(x_n, Tx_n) \to (x, y)$ in $X \times Y$. Then $x_n \to x$ and $Tx_n \to y$, so by hypothesis we must have $x \in \mathcal{D}(T)$ and $Tx = y$. Hence $(x, y) \in \mathcal{G}(T)$, proving $\mathcal{G}(T)$ is closed.

This completes the proof. ■

This property is *not* the same as the property that a *bounded* linear operator T automatically has: for a bounded linear map, the domain is all of X and it is already guaranteed to be continuous, so sequences in X with $x_n \to x$ imply $(Tx_n) \to Tx$. In contrast, an unbounded operator on a *proper* domain may fail to be closed, or may be closed but unbounded, as the next example illustrates.

Example (Differential operator)

Take $X = C[0, 1]$ (continuous functions on $[0, 1]$), and define

$$T : \mathcal{D}(T) \to X, \quad T(x) = x',$$

where $\mathcal{D}(T)$ is the set of all $x \in C[0, 1]$ that are continuously differentiable. Then T is not bounded but is indeed *closed.*

Proof. We check closedness using Theorem 4.27. Suppose $(x_n) \subset \mathcal{D}(T)$ with $x_n \to x$ in the uniform norm on $C[0, 1]$, and $(Tx_n) = x_n' \to y$ for some $y \in C[0, 1]$. Since uniform convergence on $[0, 1]$ implies $x_n' \to y$ also in the uniform norm, we can write

$$x_n(t) = x_n(0) + \int_0^t x_n'(\tau)\, d\tau.$$

Passing to the limit, we see

$$x(t) = x(0) + \int_0^t y(\tau)\, d\tau,$$

so $x'(t) = y(t)$ and $x \in \mathcal{D}(T)$. Consequently, $Tx = x' = y$, showing T is closed. ■

Observe that in this example the set $\mathcal{D}(T)$ of differentiable functions is not closed in X, so T cannot be bounded (the Closed Graph Theorem would have forced boundedness otherwise).

Closedness versus boundedness. *Closed* does not imply *bounded*, and *bounded* does not imply *closed.*

- A closed operator can be unbounded (e.g. the differentiation operator above).
- Conversely, one can have a (trivially) bounded operator that is not closed if its domain is a proper dense subspace, as in the identity on a dense but smaller domain of X. Then by extending that operator to a limit point not in the domain, one breaks the condition in Theorem 4.27.

Lemma 4.28 (Two criteria for closedness:)

Let $T : \mathcal{D}(T) \to Y$ be a bounded linear operator with $\mathcal{D}(T) \subset X$, where X and Y are normed spaces. Then:

(a) If $\mathcal{D}(T)$ is a closed subset of X, then T itself is closed.

(b) If T is closed and Y is complete, then $\mathcal{D}(T)$ is a closed subset of X.

Proof. **(a)** Suppose $(x_n) \subset \mathcal{D}(T)$ and $x_n \to x$. If $(Tx_n) \to y$ as well, we must show $x \in \mathcal{D}(T)$ and $Tx = y$. Since $\mathcal{D}(T)$ is closed and $x_n \to x$, it follows $x \in \mathcal{D}(T)$. Also, because T is bounded, $Tx_n \to Tx$ (i.e. T is continuous). Thus $y = Tx$, establishing that T is closed.

(b) Pick $x \in \overline{\mathcal{D}(T)}$, so there is a sequence $(x_n) \subset \mathcal{D}(T)$ with $x_n \to x$. By boundedness,

$$\|Tx_n - Tx_m\| \;=\; \|T\| \, \| x_n - x_m\|,$$

so (Tx_n) is Cauchy in Y. Completeness of Y ensures $Tx_n \to y$ in Y. Since T is closed, $x \in \mathcal{D}(T)$ and $Tx = y$. Hence x was already in $\mathcal{D}(T)$, showing $\mathcal{D}(T)$ is closed. ■

Bibliography

[1] John B. Conway. *A Course in Functional Analysis*. Springer, 1990. ISBN: 978-3540960423.

[2] Erwin Kreyszig. *Introductory functional analysis with applications*. Wiley, 1989. ISBN: 978-0471504597.

[3] Walter Rudin. *Functional Analysis*. McGraw-Hill, 1991. ISBN: 978-0070542365.

[4] Elias M. Stein and Rami Shakarchi. *Functional Analysis*. Princeton University Press, 2011. ISBN: 978-0691113876.