Functional Analysis

Kexing Ying

July 24, 2021

Contents

1 Introduction

We have thus far looked at abstract vector spaces in linear algebra and (metric) topological spaces in topology. In this course, we will combine these concepts and study linear metric space. In particular, we will study vector spaces equipped with a topology such that certain properties are satisfies.

In this course, we will often study the space of functions and hence the name of the course. As we have seen before, given that the codomain space possesses a certain structure, it is possible to define point-wise addition and scalar multiplications on functions, and thus, possible to equip the space with a vector space structure.

Let us recall some definitions.

Definition 1.1 (Metric)**.** A metric ρ on a non-empty set X is a function with type signature $X \times X \to \mathbb{R}^+$ such that for all $x, y, z \in X$

- $\rho(x, y) = 0 \iff x = y$;
- $\rho(x, y) = \rho(y, x)$;
- $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

Definition 1.2 (Translation Invariant)**.** A metric space (V, ρ) where V is equipped with the binary operation $(+) : V \times V \to V$ is translational invariant if for all $w, z, v \in V$,

$$\rho(w + v, z + v) = \rho(w, z).$$

Definition 1.3 (Norm)**.** A norm $\| \cdot \|$ on the vector space V (over the field $\mathbb{K}$ equipped with a modulus $| \cdot |$) is a function with type signature $X \to \mathbb{R}^+$ such that for all $x, y \in V, k \in \mathbb{K}$,

- $\|x\| = 0 \iff x = 0$;
- $\|k \cdot x\| = |k|\|x\|$;
- $\|x + y\| \leq \|x\| + \|y\|$.

We recall that a norm induces a metric by defining $\rho(x, y) = \|x - y\|$. In this case, it is possible to show that $(+)$ and $(\cdot)$ are continuous with respect to this metric and ρ is translational invariant.

Definition 1.4 (Banach Space)**.** A normed space is said to be a Banach space if it is complete, i.e. every Cauchy sequence converge.

Definition 1.5 (Separable)**.** A topological space is said to be separable if there exists a dense countable subset.

As we shall see, for $0 < 0 < \infty$, ℓ_p is separable while ℓ_∞ is not.

Definition 1.6 (Compact)**.** A topological space is said to be compact if every open cover has a finite sub-cover.

Unlike what we have seen before, as we consider infinite dimensional spaces, we will see that the Heine-Borel property will no longer hold, i.e. closed and bounded is no longer equivalent to compact.

2 Linear Spaces

Definition 2.1 (Equivalent Norms and Metrics)**.** Two norms $\|\cdot\|_k$ for $k = 1, 2$ are said to be equivalent if there exists some $M > 0$ such that for all x,

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

Similarly, two metrics ρ_k are said to be equivalent if there exists some $M > 0$ such that for all x, y,

$$\frac{1}{M}\rho_1(x, y) \leq \rho_2(x, y) \leq M\rho_1(x, y).$$

It is clear that equivalent is a symmetric relation and as we have seen before, all norms on a finite dimensional space are equivalent.

Definition 2.2 (Concave and Convex Function)**.** A function $f : V \to \mathbb{R}$ is

- concave if for all $s \in [0, 1], x, y \in V$, we have

$$sf(x) + (1 - s)f(y) \leq f(sx + (1 - s)y);$$

- convex if for all $s \in [0, 1], x, y \in V$, we have

$$sf(x) + (1 - s)f(y) \geq f(sx + (1 - s)y).$$

Proposition 2.1. If $f : \mathbb{R}^+ \to \mathbb{R}^+$ is concave and $f(0) = 0$. Then

$$f(x + y) \leq f(x) + f(y).$$

Proof. Clear by taking $s = \frac{y}{x+y}$, we have

$$(1 - s)f(x + y) = sf(0) + (1 - s)f(x + y) \leq f(s \cdot 0 + (1 - s)(x + y)) = f(x),$$

and

$$sf(x + y) = sf(x + y) + (1 - s)f(0) \leq f(s(x + y) + (1 - s) \cdot 0) = f(y).$$

Adding the two equations, we have

$$f(x + y) = (1 - s)f(x + y) + sf(x + y) \leq f(x) + f(y).$$

□

Corollary 2.1. If ρ is a metric and $\eta : \mathbb{R}^+ \to \mathbb{R}^+$ is a concave and vanishing at 0, then $\rho \circ \eta$ is also a metric.

Definition 2.3 (Linear Metric Space)**.** A vector space V over the field $\mathbb{K}$ equipped with a metric ρ on V and a metric $|\cdot - \cdot|$ on $\mathbb{K}$ is a linear metric space if $(+) : V \times V \to V$ and $(\cdot) : \mathbb{K} \times V \to V$ are continuous with respect to the induced metric.

Proposition 2.2. Any normed space is a linear metric space.

Proof. Let $(x_n, y_n) \to (x, y)$ in V^2, then we have

$$\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0.$$

Thus, $(+)$ is continuous.

Similarly, if $(\lambda_n, x_n) \to (\lambda, x)$ in $\mathbb{K} \times V$,

$$\begin{aligned}\|\lambda_n x_n - \lambda x\| &= \|\lambda_n x_n - \lambda_n x + \lambda_n x - \lambda x\| \\ &\leq \|\lambda_n x_n - \lambda_n x\| + \|\lambda_n x - \lambda x\| \\ &= |\lambda_n|\|x_n - x\| + |\lambda_n - \lambda|\|x\|.\end{aligned}$$

Now, since (λ_n) is convergent, it is bounded by some $M > 0$ and thus,

$$\|\lambda_n x_n - \lambda x\| \leq |\lambda_n|\|x_n - x\| + |\lambda_n - \lambda|\|x\| \leq M\|x_n - x\| + |\lambda_n - \lambda|\|x\| \to 0,$$

implying $(\cdot)$ is continuous. $\square$

2.1 Classical Spaces

We recall the L_p spaces from second year measure theory, and in particular, when we consider the counting measure μ, we have the nice property that

$$\int f \mathrm{d}\mu = \sum_{n=0}^{\infty} f(n),$$

and we no longer require a quotient to define the linear space as the only null-set is the empty set (thus, two function are a.e equal if and only if they are equal). In this special case, we call the resulting space ℓ_p with the p-norm

$$\|f\|_p = \left(\int |f|^p \mathrm{d}\mu\right)^{\frac{1}{p}} = \left(\sum_{n=0}^{\infty} |f(n)|^p\right)^{\frac{1}{p}}.$$

We will use the sequence notation and write $a_n := a(n)$ for $a \in \ell_p$.

As this is simply a special case of the L_p space, the inequalities proved on the L_p space remains. We will recall them here for ℓ_p spaces.

Proposition 2.3 (Hölder's Inequality)**.** Let $\frac{1}{p} + \frac{1}{q} = 1$ where $p, q \in (1, \infty)$. Then for $a = (a_i)_{i \in \mathbb{N}}, b = (b_i)_{i \in \mathbb{N}} \in \ell_p$, we have

$$|\langle a, b \rangle| \leq \|a\|_p \|b\|_q,$$

where $\langle a, b \rangle := \sum_{i \in \mathbb{N}} a_i b_i$.

Proposition 2.4 (Minkowski's Inequality)**.** Let $a, b \in \ell_p$ for some $1 \leq p \leq \infty$. Then $a + b \in \ell_p$ and

$$\|a + b\|_p \leq \|a\|_p + \|b\|_p.$$

Lemma 2.1. Let $I \subseteq \mathbb{R}$ be an interval and $\phi : I \to \mathbb{R}^+$ be a function. Then ϕ is convex if and only if for all $y \in I$, there exists a $\gamma \in \mathbb{R}$, such that for all $x \in I$,

$$\gamma(x - y) \leq \phi(x) - \phi(y).$$

Proposition 2.5 (Jensen's Inequality)**.** let $\phi \geq 0$ be convex and suppose $\sum_{i\in\mathbb{N}} \eta_i = 1$, $|\langle\alpha\rangle| < \infty$ and $\langle\phi(\alpha)\rangle$ (where $\langle\beta\rangle = \sum_{i\in\mathbb{N}} \eta_i\beta_i$), then

$$\phi(\langle\alpha\rangle) \leq \langle\phi(\alpha)\rangle.$$

Proof. By the above lemma, there exists some γ such that

$$\gamma(\alpha_j - \langle\alpha\rangle) \leq \phi(\alpha_j) - \phi(\langle\alpha\rangle).$$

Thus,

$$\begin{aligned} 0 = \sum \eta_i\gamma(\alpha_j - \langle\alpha\rangle) &\leq \sum \eta_i(\phi(\alpha_j) - \phi(\langle\alpha\rangle)) \\ &= \sum \eta_i\phi(\alpha_j) - \phi(\langle\alpha\rangle)\sum \eta_i \\ &= \langle\phi(\alpha_j)\rangle - \phi(\langle\alpha\rangle) \end{aligned}$$

where the first equality follows as γ is independent of the index i. $\square$

Proposition 2.6. For $p < p'$, $\ell_p \subseteq \ell_{p'}$. On the other hand, if $\sum \eta_i = 1$, we have $\ell_p(\eta) \supseteq \ell_{p'}(\eta)$.

Definition 2.4. Let ϕ be a convex function such that $\phi(0) = 0$, $\phi(x) \to \infty$ as $x \to \infty$. If ϕ has the doubling property such that there exists some $M > 0$ such that for all $x \in \mathbb{R}$, $\phi(2|x|) \leq M\phi(|x|)$, then

$$V := \left\{a : \mathbb{N} \to \mathbb{R} \mid \sum \eta_i\phi(|a_i|) < \infty\right\}$$

where $\sum \eta_i = 1$ is a vector space with point-wise operations.

Proposition 2.7. Given a metric space (X, ρ), there exists a complete metric space $(\tilde{X}, \tilde{\rho})$ and an isometric embedding $\iota : X \to \tilde{X}$ such that for all $x, x' \in X$, $\tilde{\rho}(\iota(x), \iota(x')) = \rho(x, x')$.

Proof. We have seen similar ideas in the completion of $\mathbb{Q}$ though completing the space with equivalence classes of mutually Cauchy sequences as elements of $\tilde{X}$. $\square$

As we have seen last year, the L_p spaces are complete, and thus are Banach spaces. Thus, we have ℓ_p spaces are also complete and are Banach spaces. In fact, the proof that ℓ_p spaces are complete is easier, in that one may show completeness through showing the point-wise limit of the sequences indeed belong to ℓ_p.

Definition 2.5. We define

- $c_0 := \{x \in \ell_\infty \mid \lim_{n\to\infty} x_n = 0\}$,
- $c := \{x \in \ell_\infty \mid \exists \lim_{n\to\infty} x_n\}$,

be subspaces of ℓ_∞.

Proposition 2.8. c_0 is complete.

Proof. Let $(x_i^n) \subseteq c_0$ be a Cauchy sequence and let $x_i = \lim_{n\to\infty} x_i^n$, then it suffices to show $\lim_{i\to\infty} x_i = 0$.

Since (x_i^n) is Cauchy, for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\|x^n - x^N\|_\infty < \frac{\epsilon}{2}.$$

Furthermore, as $x_i^N \to 0$ as $i \to \infty$, there exists some $I \in \mathbb{N}$ such that for all $i \geq I$, $|x_i^N| < \epsilon/2$. Thus, for all $i \geq I$, we have

$$\frac{\epsilon}{2} > \|x^n - x^N\|_\infty > |x_i^n - x_i^N| > |x_i^n| - \frac{\epsilon}{2} \implies \epsilon > |x_i^n|,$$

for all $n \geq N$. Hence, taking $n \to \infty$, we have

$$\epsilon > \lim_{n\to\infty} |x_i^n| = |x_i|,$$

implying $x_i \to 0$ as $i \to \infty$. □

Proposition 2.9. c is complete.

Proof. Similar to above, let (x_i^n) be Cauchy and define (x^n) such that for all $n \in \mathbb{N}$, $x^n = \lim_{i\to\infty} x_i^n$ and (x_i) such that $(x_i^n) \to (x_i)$ in ℓ_∞. It is not difficult to see that (x^n) is Cauchy and thus converges to some x.

$$\begin{tikzcd}
(x_i^n) \arrow[r, "n\to\infty"] \arrow[d, "i\to\infty"'] & (x_i) \arrow[d, "i\to\infty"] \\
(x^n) \arrow[r, "n\to\infty"] & x
\end{tikzcd}$$

Indeed, as (x_i^n) is Cauchy, for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $p, q \geq N$, $\|x_i^p - x_i^q\|_\infty < \epsilon/3$. Furthermore, there exists some $I \in \mathbb{N}$ such that for all $j \geq I$, $|x_j^p - x^p|, |x_j^q - x^q| < \epsilon/3$. Thus,

$$|x^p - x^q| < |x_j^p - x^p| + |x_j^p - x_j^q| + |x_j^q - x^q| < \frac{\epsilon}{3} + \|x_i^p - x_i^q\|_\infty + \frac{\epsilon}{3} = \epsilon.$$

Hence, it remains to show $x_i \to x$ as $i \to \infty$. Fix $\epsilon > 0$, then there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x^n - x| < \epsilon/3$. Furthermore, there exists $M \in \mathbb{N}$ such that for all $n \geq M$, $\|x_i^n - x_i\|_\infty < \epsilon/3$. Then, for all $p \geq \max\{N, M\}$, there exists $I \in \mathbb{N}$ such that for all $j \geq I$, $|x_j^p - x^p| < \epsilon/3$. Finally,

$$|x_j - x| \leq |x_j - x_j^p| + |x_j^p - x^p| + |x^p - x| \leq \|x_i - x_i^p\|_\infty + \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon,$$

implying $x_i \to x$ as $i \to \infty$ and $(x_i) \in c$. □

2.1.1 Separability

In this section we will consider the separability of different classical spaces.

Proposition 2.10. The space ℓ_p is separable for all $p \geq 1$.

Proof. It is easy to see that the subspace of all finite sequences is dense in ℓ_p. Indeed, if $x \in \ell_p$ and x^n is the sequence such that $x_i^n = x_i$ for all $i \leq n$ and $x_i^n = 0$ for all $i > n$, we have

$$\|x - x^n\|_p^p = \sum_{i=0}^{\infty} |x_i - x_i^n|^p = \sum_{i=n+1}^{\infty} |x_i|^p$$

which tends to 0 as $n \to \infty$. Now, by defining $\ell_{\mathbb{Q},n}$ as the set of rational sequences with length n, we have $d := \bigcup_{n \in \mathbb{N}} \ell_{\mathbb{Q},n}$ is dense in the space of all finite sequences. Now, since d is countable as it is a countable union of countable sets, we have found a countable dense set of ℓ_p. $\square$

Lemma 2.2. A metric space is X not separable if there exists an uncountable subset S such that for some some $k > 0$, $d(x, y) > k$ for all $x, y \in S$.

Proof. As subset of a separable *metric* space is separable, S, must be separable. But as $d(x, y) > k$ for all $x, y \in S$, no sequences of S but the constant sequence can converge. Thus, for all countable subsets of S, simply picking a point of S not in that subset suffices. $\square$

Proposition 2.11. ℓ_∞ is not separable.

Proof. Clearly the set containing all sequences of only 0 and 1s is a subset of ℓ_∞. Furthermore, this sequence is uncountable as it bijects $\mathbb{R}$ be considering the binary representation of a real number. Thus, since all distinct elements in this set have distance 1 apart, the conclusion follows by the above lemma. $\square$

Proposition 2.12. $C^k([a, b])$ (the set of k-differentiable functions from the interval $[a, b]$) for all $k \in \mathbb{N}$ is separable.

Proof. Recalling the Weierstass approximation theorem, we have for all continuous function f, there exists a sequence of polynomials f_n such that $f_n \to f$ uniformly. Thus, as the set of polynomials with rational coefficients is countable, and dense in the space of all polynomials, we have found a countable dense set of $C^0([a, b])$. Now as $C^k([a, b]) \subseteq C^0([a, b])$ for all $k \in \mathbb{N}$, we have $C^k([a, b])$ is separable as required. $\square$

Definition 2.6 (Absolutely Convergent)**.** Let X be a normed space and let (x_n) be a sequence of X, then a series $\sum_{n \in \mathbb{N}} x_n$ is said to be absolutely convergent if $\sum_{n \in \mathbb{N}} \|x_n\| < \infty$.

Proposition 2.13. A normed space X is complete if and only if for all sequences (x_n) of X such that $\sum x_n$ is absolutely convergent implies x_n converges. Thus, a normed space is Banach if and only if this property is satisfied.

Proof. See problem sheet. $\square$

By recalling the proof of the completeness of L_p spaces, we note that this property was used extensively. As a consequence, we see that $L_p(\mu)$ is separable if the measure μ is separable. In particular, the spaces $L_p(\mathbb{R}^n, \lambda)$ are separable for all n. On the other hand, L_∞ is in general not separable.

2.2 Hamel and Schauder Basis

In this small section we will introduce two new notions of basis for infinite dimensional spaces. In particular, we will introduce the Hamel basis which is a natural extension of the definition of basis for the finite dimensional case and the Schauder basis which is a notion of basis that incorporates the topological properties of the space.

Definition 2.7 (Linear Independent)**.** A set $W \subseteq V$ is linear independent if for all $(\lambda_i)_{i=1}^m \subseteq \mathbb{K}$, $(w_i)_{i=1}^n \subseteq W$,

$$\sum \lambda_i w_i = 0 \implies \lambda_i = 0$$

for all $i = 1, \cdots, n$.

Definition 2.8 (Hamel Basis)**.** A set $W \subseteq V$ is a Hamel basis for a linear space V if W is linearly independent and for all $x \in V$, there exists a unique finite linear combination of vectors $(w_i)_{i=1}^n \subseteq W$, $(\lambda_i)_{i=1}^m \subseteq \mathbb{K}$ such that

$$x = \sum \lambda_i w_i.$$

Proposition 2.14. Every linear space has a Hamel basis.

We will come back to the proof of this proposition after discussing an important result known as the Hahn-Banach theorem.

Definition 2.9 (Schauder Basis)**.** A set $W \subseteq V$ is a Schauder basis for a normed space V if W is countable, linearly independent, and for all $x \in V$, there exists a unique (possibly infinite)-sequence $(\lambda_i)_{i=1}^\infty \subseteq \mathbb{K}$ and $(w_i)_{i=1}^\infty \subseteq W$ such that

$$x = \sum_{i=1}^\infty \lambda_i w_i.$$

Proposition 2.15. If a Banach space X has a Schauder basis, then it is separable.

The proof of the above proposition is left as an exercise. It is notable that the reverse of the above is not true and was in fact a Scottish book problem which an counterexample was given by Per Enflo in 1972 who was awarded a live goose.

It is clear that for $p \geq 1$, the set $W := \{e_j \mid j \in \mathbb{N}\}$ where $(e_j)_i = \delta_{ij}$ is a Schauder basis of ℓ_p.

Recall that $c \subseteq \ell_\infty$ is the set of sequences which has a limit. By considering the sequence $x_n = 1$, we see that the set of standard basis vectors no longer form a basis of c as

$$\left\| x - \sum_{i=1}^n e_i \right\|_\infty = 1,$$

for all n. Defining $W := \{e_0 := (1, 1, \cdots)\} \cup \{e_j \mid j \in \mathbb{N}\}$, for all $(a_n) \in c$, let $\lambda_0 := \lim_{n \to \infty} a_n$ and $\lambda_n = a_n - \lambda_0$. Then, for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, $|a_n - \lambda_0| \leq \epsilon$. Thus,

$$\left\| \lambda_0 e_0 + \sum_{i=1}^n \lambda_i e_i - a \right\|_\infty = \|(0, \cdots, 0, \lambda_0 - a_n, \lambda_0 - a_{n+1}, \cdots)\|_\infty < \epsilon,$$

implying W is a Schauder basis of c.

Furthermore, one may show that $C([0,1])$ has a Schauder basis consisting of all the "spike" functions at the points $k2^{-n}$ for all $n \in \mathbb{N}$, $k = 1, \cdots 2^{-n}$.

2.3 Hilbert Spaces

Recall the definition of sesquilinear forms over some vector space $\mathbb{H}$.

Definition 2.10 (Sesquilinear Form)**.** A sesquilinear form $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{K}$ on $\mathbb{H}$ where $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$ is a function such that

- it is linear with respect to the second argument;
- and is conjugate symmetric.

We say $\langle \cdot, \cdot \rangle$ is nondegenerate if $x = 0$ if and only if $\langle x, y \rangle = 0$ for all $y \in \mathbb{H}$. Nondegenerate sesquilinear forms are called scalar products and a vector space equipped with a scalar product is called a unitary space (or a scalar product space).

We see that a scalar product induces a norm by defining $\|f\|^2 := \langle f, f \rangle$. In particular, we recall the $\ell_2, C([a,b])$ and L_2 are all scalar product spaces.

Proposition 2.16. Let $\mathbb{H}$ be a unitary space, then

- $\|f+g\|^2 + \|f-g\|^2 = 2\|f\|^2 + 2\|g\|^2$ (parallelogram identity);
- $\langle f, g \rangle = \frac{1}{4}(\|f+g\|^2 - |f-g\|^2)$ if $\mathbb{K} = \mathbb{R}$ and $\langle f, g \rangle = \frac{1}{4}\sum_{k=0,\cdots 3} i^k \|f + i^k g\|^2$ if $\mathbb{K} = \mathbb{C}$ (polarisation identity).

Definition 2.11 (Hilbert Space)**.** A unitary space is a Hilbert space if it is complete with respect to its induced norm.

Definition 2.12 (Convex Set)**.** A set S is said to be convex if for all $x, y \in S$, $(1-t)x + ty \in S$ for all $t \in [0,1]$.

Proposition 2.17 (Nearest Point Property)**.** Every nonempty closed convex set $\mathcal{K}$ in a Hilbert space $\mathbb{H}$ contains a vector of the smallest norm. Moreover, if $h \in \mathbb{H}$, there exists a unique $h_0 \in \mathcal{K}$ such that

$$\|h - h_0\| = \operatorname{dist}(h, \mathcal{K}) := \inf_{k \in K} \|h - k\|.$$

Proof. By the definition of infimum, there exists a sequence $(k_n) \subseteq \mathcal{K}$ such that

$$\lim_{n \to \infty} \|k_n\| \to d := \inf_{k \in \mathbb{K}} \|k\|.$$

Consider, for all $n, m \in \mathbb{N}$, by the parallelogram identity,

$$\left\|\frac{1}{2}(k_n - k_m)\right\|^2 = \frac{1}{2}(\|k_n\|^2 + \|k_m\|^2) - \left\|\frac{1}{2}(k_n + k_m)\right\|^2.$$

Then, as $\mathbb{K}$ is convex, $\frac{1}{2}(k_n + k_m) \in \mathbb{K}$ and so $\|(k_n + k_m)/2\|^2 \geq d^2$ and

$$\left\|\frac{1}{2}(k_n - k_m)\right\|^2 \leq \frac{1}{2}(\|k_n\|^2 + \|k_m\|^2) - d^2.$$

Thus, taking $n, m \to \infty$, the right hand side tends to zero and so (k_n) is Cauchy and hence convergent. Now as $\mathscr{K}$ is closed, the limit is in $\mathscr{K}$ and hence the statement. $\square$

Definition 2.13 (Orthogonal Systems)**.** Let $(\mathbb{H}, \langle\cdot,\cdot\rangle)$ be an unitary space. A set of vectors $\{e_j \in \mathbb{H} \mid j \in J\}$ for some index set J is called an orthogonal system if for all distinct $i, j \in J$,

$$\langle e_i, e_j\rangle = 0.$$

If furthermore, $\langle e_i, e_i\rangle = 1$ for all $i \in J$, then we say the system is orthonormal.

Proposition 2.18. Every orthogonal system is linearly independent.

Proof. Exercise. $\square$

Using Zorn's lemma, one can show that every unitary space contains an orthogonal basis.

Definition 2.14 (Fourier Coefficients)**.** Let $(\mathbb{H}, \langle\cdot,\cdot\rangle)$ be an unitary space and let $\{e_j \in \mathbb{H} \mid j \in J\}$ be an orthogonal system. Then for each $f \in \mathbb{H}$, the Fourier coefficients with respect to the orthogonal system are

$$c_j := \frac{\langle e_j, f\rangle}{\|e_j\|^2}.$$

Proposition 2.19. If $f = \sum_{k\in\mathbb{N}} \alpha_k e_k$, then $a_j = c_j$.

Proof. Let $n < m$ and $S_m := \sum_{k=1}^m \alpha_k e_k$. Then, $\langle S_m, e_n\rangle = \overline{\alpha_n}\|e_n\|^2$. Thus, we have

$$|\overline{\alpha_n}\|e_n\|^2 - \langle f, e_n\rangle| = |\langle S_m, e_n\rangle - \langle f, e_n\rangle| = |\langle S_m - f, e_n\rangle| \le \|S_m - f\|\|e_n\| \to 0$$

as $m \to \infty$. $\square$

Proposition 2.20. Suppose $(e_j)_{j\in\mathbb{N}}$ is orthogonal and $g(a_1, \cdots, a_n) := \left\|f - \sum_{j=1}^n a_j e_j\right\|^2$. Then g attains its minimum at $a_j = c_j$. Furthermore, we have

$$\sum_{j=1}^{\infty} |c_j|^2 \|e_k\|^2 \le \|f\|^2.$$

This inequality is known as Bessel's inequality.

Proof. Exercise. $\square$

Definition 2.15. An orthogonal system $\{e_j \mid j \in J\}$ is called complete if for all $f \in \mathbb{H}$, if $\langle f, e_j\rangle = 0$ for all j, then $f = 0$.

Proposition 2.21. The following are equivalent

1. $(e_j)_{j\in\mathbb{N}}$ is complete;
2. $\left\|f - \sum_{j=1}^n c_j e_j\right\| \to 0$ as $k \to \infty$ where c_j are the Fourier coefficients;
3. $\|f\|^2 = \sum_{j=1}^{\infty} |c_j|^2 \|e_k\|^2$ for all $f \in \mathbb{H}$.

Proof. Exercise. $\square$

2.4 Finite Dimensional Spaces

Theorem 1. Let $(X, \|\cdot\|)$ be a finite dimensional normed space and let $\{e_i\}$ be a basis for X. Then, there exists some $M, m \in \mathbb{R}^+$ such that for all $x = \sum_{i=1}^n a_i e_i$,

$$m \sum_{i=1}^n |a_i| \leq \|x\| \leq M \sum_{i=1}^n |a_i|.$$

Proof. WLOG. assume $\sum_{i=1}^n |a_i| = 1$ since if otherwise, we may just scale x by $1/\sum |a_i|$. Then the function $f : (a_i) \mapsto \|\sum_i a_i e_i\|$ for all sequences (a_i). Now, by checking that f is continuous, and by considering that the set $\{(a_i) \mid \sum \|a_i = 1\|\}$ is a closed subspace, f must attain a minimum m implying the left hand side inequality. On the other hand, consider the inequality

$$\left\|\sum_{i=1} a_i e_i\right\| \leq \sum |a_i| \|e_j\| \leq \max_k \|e_k\| \sum |a_i| = \max_k \|e_k\|.$$

Hence, it suffices to choose $M = \max \|e_k\|$. □

A direct corollary is that all norms on finite dimensional spaces are equivalent.

Definition 2.16 (Equivalent Norms)**.** Two norms $\|\cdot\|_1, \|\cdot\|_2$ on X are said to be equivalent if there exists some $C \in \mathbb{R}^+$ such that

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

for all $x \in X$.

Proposition 2.22. Equivalence of norms is an equivalence relation.

Proof. Easy check. □

Corollary 2.2. Every norm on a finite dimensional space is equivalent.

Proof. If $\|\cdot\|_1, \|\cdot\|_2$ are two norms on X, then let $\{e_i\}$ be a basis of X. Thus, by simply defining the norm

$$\|x\| = \left\|\sum_{i=1}^n a_i e_i\right\| := \sum_i |a_i|,$$

we have $\|\cdot\|_1$ is equivalent to $\|\cdot\|$ which is equivalent to $\|\cdot\|_2$. Hence, $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$ by transitivity. □

Proposition 2.23. Every finite dimensional space over a complete field is complete.

Proof. Follows by considering the individual coefficients of a Cauchy sequence with respect to some basis is also Cauchy. □

Proposition 2.24. Every compact set of a normed space is closed and bounded.

Proof. As a normed space is a metric space, it suffices to consider sequential compactness. In particular, if a set is not closed, then it contains a sequence converging to a point not within the set. Then every sub-sequence of that sequence also converges to that point outside of the set, and so is not compact. On the other hand, if the set is not bounded, we can construct a sequence with norms tending to ∞. It is then clear that the sequence does not contain any convergent subsequence. □

We note that the converse of the above proposition is not true for infinite dimensional spaces. Indeed, we see that the canonical basis of ℓ_2 is closed and bounded but not complete.

Proposition 2.25. If $(X, \|\cdot\|)$ is a finite dimensional normed space, then every closed and bounded set is compact (this property is known as the Heine-Borel property).

Proof. Follows by choosing a subsequence for each coefficient by using Bolzano-Weierstrass. □

In fact this property is sufficient to determine whether or not a normed space is finite dimensional.

Lemma 2.3 (Riesz's lemma)**.** Let Y be a closed proper subspace of a subspace Z of X where $(X, \|\cdot\|)$ is a normed space. Then for any $\theta \in (0,1)$, there exists some $z \in Z$ such that

$$\|z\| = 1 \text{ and } \|y - z\| \geq \theta$$

for all $y \in Y$.

Proof. Let $v \in Z \setminus Y$, and define $a := \inf_{y \in Y} \|y - v\| > 0$. Then, as Y is closed, $\|y_0 - v\|$ attains a for some $y_0 \in Y$. Thus,

$$0 < a = \|v - y_0\| \leq \frac{a}{\theta}.$$

Then, defining

$$z := \frac{v - y_0}{\|v - y_0\|} =: c(v - y_0),$$

we have for all $y \in Y$,

$$\|z - y\| = \|c(v - y_0) - y\| = c\left\|v - \left(y_0 + \frac{1}{c}y\right)\right\|.$$

Since $y_0 + \frac{1}{c}y =: y_1 \in Y$, we have $\|v - y_1\| \geq a$, and so

$$\|z - y\| = c\|v - y_1\| \geq ca = \frac{a}{\|v - y_0\|} \geq \frac{a}{a/\theta} = \theta.$$

□

Proposition 2.26. A normed space is finite dimensional if and only if $\{x \mid \|x\| = 1\}$ is compact.

Proof. Clearly, if X is finite dimensional, then as $\{x \mid \|x\| = 1\}$ is closed and bounded, it is compact.

Let $(X, \|\cdot\|)$ be an infinite dimensional normed space and suppose $\{e_i\}_{i=1}^{\infty}$ is a countable infinite linearly independent subset of X. Define $Z_n = \mathrm{span}\{e_i \mid i = 1, \cdots, n\}$, and for each n, by Riesz's lemma, define $z_n \in Z_{n+1}$ such that $\|z_n\| = 1$ and $\|z_n - z_i\| \geq 1/2$ for all $i = 1, \cdots, n$. Thus, we have defined a sequence in $\{x \mid \|x\| = 1\}$ which does not contain any convergent subsequence, and hence, $\{x \mid \|x\| = 1\}$ is not compact. $\square$

3 Linear Operators

3.1 Bounded Linear Operators

Definition 3.1 (Bounded Linear Operator)**.** A linear map $T : X \to Y$ between normed spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ is a bounded linear operator if there exists some $c \in \mathbb{R}^+$ such that,

$$\|Tx\|_Y \leq C\|x\|_X$$

for all $x \in X$. In this case, one may define the operator norm by,

$$\|T\| := \sup_{x \in X; x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X}.$$

Proposition 3.1. Every linear map from a finite dimensional normed spaces is bounded.

Proof. Simply bound T by the maximum of the norm of the basis. $\square$

Proposition 3.2. Let X, Y be Banach spaces and let $T : X \to Y$ be a continuous linear map, then for all compact $\mathcal{K} \subseteq X$, $T(\mathcal{K}) \subseteq Y$ is also compact.

Proof. Follows by recalling that the continuous image of a compact space is compact (by pull-back of the open-cover). $\square$

Proposition 3.3. If $T : X \to Y$ is a linear map between normed spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$, then the following are equivalent,

1. T is continuous;
2. T is Lipschitz continuous;
3. T is bounded;
4. T is continuous at some point $x_0 \in X$.

Proof. Clearly (3) $\implies$ (2) $\implies$ (1) $\implies$ (4), so let us first show (4) $\implies$ (1)). If T is continuous at x_0, then for all $\epsilon > 0$, there exists some $\delta > 0$ such that for all $x \in B_\delta(x_0)$, $\epsilon > \|Tx - Tx_0\|$. Then, for all $\tilde{x} \in X$, we have

$$\epsilon > \|Tx - Tx_0\| = \|Tx + T(\tilde{x} - x_0) - Tx_0 - T(\tilde{x} - x_0)\| = \|T(\tilde{x} + (x - x_0)) - T\tilde{x}\|.$$

Thus, for all $x \in B_\delta(\tilde{x})$, we have $x - \tilde{x} + x_0 \in B_\delta(x_0)$ and hence,

$$\epsilon > \|T(\tilde{x} + (x - \tilde{x} + x_0 - x_0)) - T\tilde{x}\| = \|Tx - T\tilde{x}\|,$$

which implies T is continuous.

Now, it suffices to show that (1) $\implies$ (3). Since T is continuous at 0, there exists some $\delta > 0$ such that for all $x \in B_\delta(0)$, $\|Tx\| < 1$. Then for all $x \in X$, we have

$$\left\|\frac{\delta x}{2\|x\|}\right\| = \frac{\delta}{2} < \delta,$$

and so,

$$1 > \left\| T \left\| \frac{\delta x}{2\|x\|} \right\| \right\| = \frac{\delta}{2\|x\|} \|Tx\|,$$

implying $Tx < 2\|x\|/\delta$ for all x. ∎

Definition 3.2. We denote the space of linear bounded operators between two normed spaces X, Y equipped with the operator norm by $\mathcal{L}(X, Y)$.

Proposition 3.4. Let $(X, \|\cdot\|_X)$ be a normed space and $(Y, \|\cdot\|_Y)$ be a Banach space. Then $\mathcal{L}(X, Y)$ is also a Banach space.

Proof. Exercise. ∎

We will now recall the Banach contraction mapping theorem.

Definition 3.3 (Contraction)**.** A map $T : X \to X$ on a metric space (X, ρ) is a contraction if there exists some α, $0 < \alpha \leq 1$ such that for all $x, y \in X$,

$$\rho(Tx, Ty) \leq \alpha \rho(x, y).$$

T is a strict contraction if $a < 1$.

Definition 3.4 (Fixed Point)**.** A point $x \in X$ is a fixed point of the map $T : X \to X$ if $Tx = x$.

Theorem 2 (Banach Contraction Mapping)**.** If X is a complete metric space and $T : X \to X$ is a strict contraction, then T has a unique fixed point.

Proof. Exercise/see second year analysis. ∎

3.2 Dual Space and Dual Operators

3.2.1 Dual Space

Definition 3.5 (Dual Space)**.** Given a normed space $(X, \|\cdot\|)$, its dual space is $X^* := \mathcal{L}(X, \mathbb{K})$. The elements of the dual spaces are called continuous functionals on X.

Proposition 3.5. $(\ell_p)^* = \left\{T : \ell_p :\to \mathbb{K} \mid T(x) = \sum_i y_i x_i, y \in \ell_q\right\}$ where $1/p + 1/q = 1$ and $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Proof. Let $T(x) = \sum_i y_i x_i$ for some $y \in \ell_q$, it is clear that T is linear. Consider, by Hölder's inequality

$$\|T(x)\| = \left| \sum_i y_i x_i \right| \leq \|x\|_p \|y\|_q.$$

Thus, $\|T\| \leq \|y\|_q < \infty$, implying $T \in (\ell_p)^*$.

Now, let $\phi \in (\ell_p)^*$. Fixing $y_n = \phi(e_n)$, it suffices to show that $(y_n) \in \ell_q$ since it is clear that $\phi(x) = \sum x_i y_i$ by linearity. Setting $x_i = |y_i|^q / y_i$ for all $y_i \neq 0$ and $x_i = 0$ for $y_i = 0$, consider that $\phi(\sum x_i e_i) = \sum x_i y_i = \sum |y_i|^q$. On the other hand,

$$|\phi(x)| \leq \|\phi\| \|x\|_p = \|f\| \left(\sum |x_i|^p\right)^{1/p} = \|\phi\| \left(\sum \frac{|y_i|^{qp}}{|y_i|^p}\right)^{1/p}.$$

As $1/p + 1/q = 1$, we have $qp = p + q$ and so,

$$|\phi(x)| \leq \|\phi\| \left(\sum \frac{|y_i|^{qp}}{|y_i|^p} \right)^{1/p.} = \|\phi\| \left(\sum |y_i|^q \right)^{1/p}.$$

Thus,

$$\sum |y_i|^q \leq \|\phi\| \left(\sum |y_i|^q \right)^{1/p},$$

implying

$$\|y_i\|_q = \left(\sum |y_i|^q \right)^{1/q} = \left(\sum |y_i|^q \right)^{1-1/p} = \frac{\sum |y_i|^q}{\left(\sum |y_i|^q \right)^{1/p}} \leq \|\phi\| < \infty,$$

and $(y_i) \in \ell_q$. $\square$

Proposition 3.6. If X is a Hilbert space, then for all $y \in X$,

$$T_y(x) := \langle y, x \rangle$$

is a bounded linear functional on X.

Proof. Follows by the Cauchy-Schwarz inequality. $\square$

In year two linear algebra we saw that all linear functionals on a finite dimensional space are of the form as above. In fact, this is true for all Hilbert spaces.

Lemma 3.1. Let X be a Hilbert space and Y a closed subspace of X. Then $X = Y \oplus Y^{\perp}$.

Proof. Y is complete since it is a closed subspace of a complete space. As $Y \cap Y^{\perp} = \{0\}$, it suffices to show that $Y + Y^{\perp} = X$. Let $x \in X$, then as subspaces are convex, by the nearest point property, there exists some $y \in Y$ such that $\|y - x\| = \operatorname{dist}(x, Y)$. Defining $y' = x - y$, we have $x = y + y'$ and it remains to show $y' \in Y^{\perp}$.

Suppose otherwise, there exists some $\tilde{y} \in Y$ such that $\langle \tilde{y}, y' \rangle > 0$ (we may assume the by simply multiplying by a scalar). Then, for $t > 0$,

$$\begin{aligned}
\|(y + t\tilde{y}) - x\|^2 &= \langle y + t\tilde{y} - x, y + t\tilde{y} - x \rangle \\
&= \langle y - x, y - x \rangle + \langle t\tilde{y}, y - x \rangle + \langle y - x, t\tilde{y} \rangle + t^2 \langle \tilde{y}, \tilde{y} \rangle, \\
&= \operatorname{dist}(x, Y)^2 - 2t\langle \tilde{y}, x \rangle + t^2 \|\tilde{y}\|^2
\end{aligned}$$

were for small enough t, $2t\langle \tilde{y}, x \rangle > t^2 \|\tilde{y}\|^2$ and so, $\|(y + t\tilde{y}) - x\| < \operatorname{dist}(x, Y)$, contradiction! **N.B. I am unclear about the last equality!** $\square$

Theorem 3 (Riesz Representation Theorem)**.** Let X be a Hilbert space and $l \in X^*$. Then there exists a unique $z \in X$ such that for all $x \in X$, $l(x) = \langle z, x \rangle$.

Proof. Uniqueness is easy to check since, if $\langle z, \cdot \rangle = \langle z', \cdot \rangle$, then $\langle z - z', \cdot \rangle = 0$ implying $z - z' = 0$.

If $l = 0$ then $z = 0$ suffices so suppose otherwise, i.e. $N := \ker l \neq X$. Then, as $N = l^{-1}(\{0\})$ where $\{0\}$ is closed, N is also closed, and thus, by the above lemma $N \oplus N^{\perp} = X$ and so, there exists some $z_0 \in N^{\perp}$ $z_0 \neq 0$. Define

$$v = l(x)z_0 - l(z_0)x.$$

We see that $l(v) = l(x)l(z_0) - l(z_0)l(x) = 0$ so $v \in N$, and hence $\langle z_0, v\rangle = 0$. Unfolding the definition, we have

$$0 = \langle z_0, v\rangle = \langle z_0, l(x)z_0 - l(z_0)x\rangle = l(x)\|z_0\|^2 - l(z_0)\langle z_0, x\rangle,$$

implying

$$l(x) = \frac{l(z_0)\langle z_0, x\rangle}{\|z_0\|^2} = \left\langle \frac{l(z_0)}{\|z_0\|^2} z_0, x \right\rangle.$$

Hence, picking $z = (l(z_0)/\|z_0\|^2)z_0$ suffices. $\square$

With this, is easy to check that the map $\phi : X \to X^* : v \mapsto \langle \cdot, v\rangle$ is in fact a isometric anti-isomorphism.

Let us now consider the dual space of ℓ_p for $p \in [1, \infty]$. In the case that $p = 2$, we know that ℓ_2 is a Hibert space, and so the Riesz representation theorem provides an isomorphism $\ell_2 \cong (\ell_2)^*$. In the case that $p \in (1, \infty)$, as we have seen every linear functional of ℓ_p is of the form

$$T : \ell_p \to \mathbb{K} : x \mapsto \sum_i x_i y_i,$$

for some $y \in \ell_q$, $1/p + 1/q = 1$, and thus, by checking linearity, we fine $(\ell_p)^* \cong \ell_q$.

In the case the $p = 1$, for all $y \in \ell_\infty$, $f_y : \ell_1 \to \mathbb{K}$ defined to be $f_y(x) = \sum_i x_i y_i$ is a continuous linear functional since

$$|f_y(x)| = \left|\sum_i x_i y_i\right| \leq \|y\|_\infty \left|\sum_i x_i\right| = \|y\|_\infty \|x\|_1 < \infty.$$

In particular, this provides an injection $\ell_\infty \hookrightarrow (\ell_1)^*$. By the same argument, we can show $\ell_1 \hookrightarrow (\ell_\infty)^*$.

3.2.2 Dual Operators

Definition 3.6 (Dual Operator)**.** Given a Hilbert space X, and $L \in \mathcal{L}(X, X)$, we define the dual operator $L^* : X^* \to X^*$ by

$$L^* f := f \circ L.$$

We equip L^* with the operator norm

$$\|L^*\| = \sup_{f \neq 0} \frac{\|L^* f\|}{\|f\|}.$$

Proposition 3.7. $\|L^*\| = \|L\|$ and the following properties hold,

- $(S+T)^* = S^* + T^*$;
- for all $\alpha \in \mathbb{K}, (\alpha T)^* = \overline{\alpha} T^*$;
- $(T^*)^* = T$;
- $(ST)^* = T^* S^*$.

Consider X a Hilbert space and let $T \in \mathcal{L}(X, X)$. Then by Riesz representation, for all $f \in X^*$, there exists some $z_{T^*f} \in X$ such that $T^*f(x) = \langle z_{T^*f}, x\rangle$. On the other hand, as $T^*f(x) = f(Tx)$, by Riesz, there exists some z_f such that $T^*f(x) = f(Tx) = \langle z_f, Tx\rangle$. Comparing both equations, we have

$$\langle z_{T*f}, x\rangle = \langle z_f, Tx\rangle = \langle T^\dagger z_f, x\rangle,$$

where $T^\dagger$ is the adjoint of T. Thus, $T^\dagger z_f = z_{T^*f}$.

Definition 3.7 (Dual Operator again)**.** If X_1, X_2 are Hilbert spaces and $T \in \mathcal{L}(X_1, X_2)$, the dual operator of T is

$$T^* : X_2^* \to X_1^* : f \mapsto f \circ T.$$

By Riesz representation, $X_i^* \cong X_i$, and we may view T^* as a linear operator on $X_2 \to X_1$. In particular, for all $f \in X_2^*$, $T^*f \in X_1^*$ so there exists some $z_{T^*f} \in X_1$ such that $T^*f(x) = \langle z_{T^*f}, x\rangle$. On the other hand, $T^*f(x) = f(Tx)$ and so, there exists some $z_f \in X_2$ such that $T^*f(x) = \langle z_f, Tx\rangle$. So,

$$\langle z_f, Tx\rangle = \langle z_{T^*f}, x\rangle.$$

Then, we may define $T^\dagger : X_2 \to X_1$ such that $T^\dagger z_f = z_{T^*f}$.

Consider $\nabla : D(\nabla) \to L^2$ where $D(\nabla) \subseteq L^2$. Suppose for now that $D(\nabla) = C_0^\infty$, i.e. smooth, compactly supported functions. As demonstrated above, one may find the dual operator ∇^* of ∇ such that

$$\langle g, \nabla\phi\rangle = \langle \nabla^* g, \phi\rangle$$

for all $g \in L^2$, $\phi \in C_0^\infty$. This operator is known as the weak derivative and we shall treat this topic properly later in this course

Proposition 3.8. Given a linear operator $T \in \mathcal{L}(X_2, X_1)$ where X_1, X_2 are normed spaces, the map

$$h : X_1 \times X_2 \to \mathbb{K} : (x, y) \in X_1 \times X_2 \mapsto \langle x, Ty\rangle$$

is a sesquilinear form.

Definition 3.8 (Bounded Sesquilinear Form)**.** A sesquilinear form h on a normed space is bounded if

$$\|h(\cdot,\cdot)\| := \sup_{x,y\neq 0} \frac{|h(x,y)|}{\|x\|\|y\|} < \infty.$$

Proposition 3.9. Let X_1, X_2 be Hilbert spaces and let h be a bounded sesquilinear form. Then h has the representation

$$h(x, y) = \langle Sx, y\rangle,$$

for some $S \in \mathcal{L}(X_1, X_2)$ such that $\|S\| = \|h\|$.

Proof. For all $x \in X_1$, as $h(x, \cdot) : X_2 \to \mathbb{K} \in X_2^*$, by Riesz, there exists some z_x such that $h(x, \cdot) = \langle z_x, \cdot \rangle$. Thus, simply defining $S : X_1 \to X_2 : x \mapsto z_x$ suffices after simple checks for linearity.

Consider

$$\|h\| = \sup_{x,y \neq 0} \frac{|h(x,y)|}{\|x\|\|y\|} = \sup_{x,y \neq 0} \frac{|\langle Sx, y \rangle|}{\|x\|\|y\|} \geq \sup_{x \neq 0} \frac{|\langle Sx, Sx \rangle|}{\|x\|\|Sx\|} = \sup_{x \neq 0} \frac{\|Sx\|}{\|x\|} = \|S\|.$$

On the other hand,

$$\|h\| = \sup_{x,y \neq 0} \frac{|\langle Sx, y \rangle|}{\|x\|\|y\|} \leq \sup_{x,y \neq 0} \frac{\|Sx\|\|y\|}{\|x\|\|y\|} = \|S\|,$$

where the inequality is due to Cauchy-Schwarz. Thus $\|h\| = \|S\|$ as required. □

3.3 Tietze-Urysohn and Hahn-Banach

Lemma 3.2 (Urysohn's Lemma)**.** For $A, B \subseteq X$ closed with $A \cap B$, then there exists a continuous function such that $h \mid A = 1$ and $h \mid_B = 0$.

Proof. Consider

$$h(x) := \frac{\operatorname{dist}(x, B)}{\operatorname{dist}(x, A) + \operatorname{dist}(x, B)}.$$

□

Theorem 4 (Tietze-Urysohn)**.** If $f : A \to \mathbb{R}$ is a continuous function where A is a closed set in a metric space X, then there exists a continuous function $\tilde{f} : X \to \mathbb{R}$ such that $\tilde{f} \mid_A = f$ and $\|\tilde{f}\| = \|f\|$.

Proof. We may assume $\|f\| = 1$ and $0 \leq f \leq 1$ by scaling the function.

Define $f_0 = f$ and $f_{n+1} = f_n - g_n \mid_A$ where g_n is defined using Urysohn's Lemma such that

$$g_n(x) = \begin{cases} 0, & x \in f_n^{-1}([0, \frac{1}{3}(\frac{2}{3})^n]); \\ \frac{1}{3}\left(\frac{2}{3}\right)^n, & x \in f_n^{-1}([(\frac{2}{3})^n, (\frac{2}{3})^{n+1}]). \end{cases}$$

In some sense, what we are doing taking g_{n+1} to approximate the errors of $f_n - g_n \mid A$.

By construction, we see that $0 \leq f_n \leq (\frac{2}{3})^n$, and $\sum_{i=0}^n g_i = f - f_{n+1}$, defining $\tilde{f} = \sum_i g_i$ (which exists as $0 \leq g_n \leq \frac{1}{3}(\frac{2}{3})^n$, implying $\sum^n g_i$ is Cauchy, hence convergent by the completeness of $C(A)$), we have

$$\sum_{i=0}^n g_i \mid_C - f = -f_{n+1} \to 0,$$

and so $\tilde{f} \mid_C = f$ as required. □

Definition 3.9. A functional $p : X \to \mathbb{R}$ is called sublinear if

- $p(x+y) \leq p(x) + p(y)$;
- for all $\alpha \geq 0$, $p(\alpha x) = \alpha p(x)$.

Theorem 5 (Hahn-Banach)**.** Let X be a normed space, Z a proper subspace of X, $p : X \to \mathbb{R}$ a sublinear functional and $f : Z \to \mathbb{R}$ such that $f(x) \leq p(x)$ for all $x \in Z$. Then, there exists some linear functional $\tilde{f} : X \to \mathbb{R}$ such that

$$\tilde{f}\mid_Z = f \text{ and } \tilde{f}(x) \leq p(x)$$

for all $x \in X$.

Proof. Let us first consider the special case of extending f by 1-dimension. Let $v \in X \setminus Z$ and $W := \operatorname{span}(Z, v)$. Define

$$\tilde{f} : W \to \mathbb{R} : z + \lambda v \mapsto f(z) + \lambda\alpha$$

for some $\alpha \in \mathbb{R}$. It suffices to show that

$$\tilde{f}(z + \lambda v) \leq p(z + \lambda v).$$

In particular, if $\lambda > 0$, then $\alpha \leq p\left(\frac{1}{\lambda}z + v\right) - \tilde{f}\left(\frac{1}{\lambda}z\right)$, and if $\lambda < 0$, then $\alpha \geq -p\left(\frac{1}{-\lambda}z - v\right) + \tilde{f}\left(\frac{1}{-\lambda}z\right)$. Combining the two, we have for all $\lambda > 0$,

$$-p\left(\frac{1}{\lambda}z - v\right) + \tilde{f}\left(\frac{1}{\lambda}z\right) \leq \alpha \leq p\left(\frac{1}{\lambda}z + v\right) - \tilde{f}\left(\frac{1}{\lambda}z\right),$$

and by rescaling, the condition becomes

$$-p\left(z - v\right) + \tilde{f}\left(z\right) \leq \alpha \leq p\left(z + v\right) - \tilde{f}\left(z\right).$$

This holds if and only if

$$-p(z_1 - v) + \tilde{f}(z_1) \leq p(z_2 + v) + \tilde{f}(z_2)$$

for all $z_1, z_2 \in Z$. Indeed, this holds as,

$$\tilde{f}(z_1) + \tilde{f}(z_2) = \tilde{f}(z_1 + z_2) \leq p(z_1 + z_2) = p((z_1 - v) + (z_2 + v)) \leq p(z_1 - v) + p(z_2 + v).$$

With this 1-dimensional extension in mind, the result follows by Zorn's lemma. In particular, given two extensions of f, $(M_1, g_1), (M_2, g_2)$ where $Z \subseteq M_1, Z \subseteq M_2$ and $g_1\mid_Z = g_2\mid_Z = f$, we define the partial order $\prec$ such that $(M_1, g_1) \prec (M_2, g_2)$ if and only if $M_1 \subseteq M_2$ and $g_2\mid_{M_1} = g_1$. Then, for any chain of extensions $(M_\gamma, g_\gamma)_{\gamma \in T}$, define $M = \bigcup_{\gamma \in T} M_\gamma$ and

$$G : M \to \mathbb{R} : x \mapsto g_\gamma(x) \text{ if } x \in M_\gamma.$$

It is clear that G is continuous and for all (M_γ, g_γ), we have $(M_\gamma, g_\gamma) \prec (M, G)$. Hence, every chain of extensions has an upperbounds, and thus, by Zorn's lemma, there is an maximal extension (M, G). Finally, it suffices to show that $M = X$ which is clear since if otherwise, we may extend G with an additional dimension, contradicting the maximality of (M, G). □

3.4 Application of Hahn-Banach

In this section we will demonstrate some applications of the Hahn-Banach theorem.

Corollary 3.1 (Existence of tangent functional)**.** Let X be a normed space and let $x \in X$, then there exists some $l \in X^*$ such that $\|l\| = 1$ and $l(x) = \|x\|$.

Proof. Apply Hahn-Banach to $f : \mathbb{R} \cdot x \to \mathbb{R} : \lambda \mapsto \lambda \|x\|$. ☐

Corollary 3.2. Let X be a normed space, for all $x \in X$,

$$\|x\| = \sup\{|l(x)| \mid l \in X^*, \|l\| = 1\}.$$

Proof. For all $\|l\| = 1$, we have $|l(x)| \leq \|l\|\|x\| = \|x\|$. Now, by the previous corollary, $\|x\|$ is achieved by some $l \in X^*, \|l\| = 1$ and so, $\|x\| = \sup\{|l(x)| \mid l \in X^*, \|l\| = 1\}$ as required. ☐

Theorem 6 (Banach Limit Theorem)**.** There exists $L \in (\ell_\infty)^*$ such that

- $\|L\| = 1$;
- if $x \in c \subseteq \ell_\infty$, then $L(x) = \lim_{n \to \infty} x_n$;
- if $x \in \ell_\infty$ and $x_i \geq 0$ for all i, then $L(x) \geq 0$;
- if $x \in \ell_\infty$ and $x_n' := x_{n+1}$, then $L(x) = L(x')$.

Theorem 7 (Müntz-Szäsz)**.** Suppose $0 < \lambda_1 < \lambda_2 < \cdots$, and let X be the closure of the set of linear combinations of t^{λ_j} in $\mathcal{C}[0, 1]$. Then

- if $\sum_j \frac{1}{\lambda_j} = \infty$, then $X = \mathcal{C}[0, 1]$;
- if $\sum_j \frac{1}{\lambda_k} < \infty$, then for $\lambda \notin \{\lambda_j\}_{j \in \mathbb{N}}, t^\lambda \notin X$, and so $X \subsetneq \mathcal{C}[0, 1]$.

Proposition 3.10. Let Y be a closed proper subspace of a normed space X. Then, there exists some $f \in X^*$ such that $\|f\| = 1$ and $f(y) = 0$ for all $y \in Y$.

Proof. Let $x \in X \setminus Y$. If $X = \text{span}(\{x\} \cup Y)$, then the map $\lambda x + y \mapsto \lambda$ suffices. On the other hand, defining $l : \lambda x + y \mapsto \lambda$, with Hahn-Banach, there exists some $f \in X^*$ such that $f \mid_Y = l = 0$ and $\|f\| = \|l\| = 1$. ☐

Proposition 3.11. If X^* is separable, then so is X.

Proof. Let $S^* := \{\phi \in X^* \mid \|\phi\| = 1\}$. Then, as a subset of a separable space is separable, S^* is separable by some $\{\phi_n\} \subseteq S^*$. By definition, for each n, we may define x_n such that $\|x_n\| = 1$ and $\phi_n(x_n) > 1/2$. Now, let

$$\mathcal{D} := \overline{\text{span}\{x_n \mid n \in \mathbb{N}\}}.$$

By definition, it is clear that $\mathcal{D}$ is dense by considering the density of $\mathbb{Q}$ in $\mathbb{R}$ and thus, it suffices to show $\mathcal{D} = X$.

Suppose $\mathcal{D} \neq X$, then by the previous proposition, there exists a linear functional $\phi \in S^*$ such that $\phi(y) = 0$ for all $y \in \mathcal{D}$. Now, as $\{\phi_n\}$ is dense in S^*, there exists n such that $\|\phi - \phi_n\| < 1/2$. Hence,

$$\begin{aligned}|\phi(x_n)| &\geq |\phi_n(x_n)| - |\phi_n(x_n) - \phi(x_n)| > 1/2 - |(\phi_n - \phi)(x_n)| \\ &\geq 1/2 - \|\phi - \phi_n\|\|x_n\| > 1/2 - 1/2 = 0.\end{aligned}$$

But $x_n \in \mathcal{D}$ implying $\phi(x_n) = 0$, contradiction! □

3.4.1 Stieltjes Integral

Stieltjes integral is a method of defining integration similar to that of the Riemann integral (one may construct a similar version to the Lebesgue integral as well). In particular, given some function ω, we replace the individual segments $x_{i+1} - x_i$ with $|\omega(x_{i+1}) - \omega(x_i)|$.

Definition 3.10 (Total Variation of Functions)**.** The total variation of a function ω on the interval $[a, b]$ is

$$V(\omega) := \sup_{p \in \mathcal{P}} \sum_{i=1}^{|p|-1} |\omega(x_{i+1}) - \omega(x_i)|,$$

where $\mathcal{P}$ is the set of all partitions on $[a, b]$.

In the case that the total variation is finite, we may define the lower Stieltjes sum of the function f to be

$$s_\omega(p) := \sum_{i=1}^{|p|-1} f(\underline{x_i})(\omega(x_{i+1}) - \omega(x_i)),$$

and the upper sum to be

$$S_\omega(p) := \sum_{i=1}^{|p|-1} f(\overline{x_i})(\omega(x_{i+1}) - \omega(x_i)),$$

where p is a partition, $f(\underline{x_i}) = \min_{x \in [x_i, x_{i+1}]} f(x)$ and $f(\overline{x_i}) = \max_{x \in [x_i, x_{i+1}]} f(x)$. We observe that $|s_\omega|, |S_\omega| \leq \|f\|_\infty V(\omega)$.

If $\inf S_\omega = \sup s_\omega$ as the size of the segments of the partition tends to zero, then we define the Stieltjes integral of f with respect to $\mathrm{d}\omega$ to be

$$\int_a^b f \mathrm{d}\omega := \inf S_\omega.$$

In the case that ω is differentiable then we see $\int f \mathrm{d}\omega = \int f \omega' \mathrm{d}x$.

Proposition 3.12. Let $l \in \mathcal{C}([a, b])^*$ can be represented as

$$l(f) = l_\omega(f) := \int f \mathrm{d}\omega,$$

for some ω with finite total variation and $\|l_\omega\| = V(\omega)$.

Proof. By the Hahn-Banach theorem, as $\mathcal{C}([a,b]) \subseteq \mathcal{B}([a,b])$, there exists some $\tilde{l} \in \mathcal{B}([a,b])^*$ such that $\|l\| = \|\tilde{l}\|$ and $\tilde{l}\mid_{\mathcal{C}(a,b)} = l$. Define $\omega(t) := \tilde{l}(\chi_{[a,t]})$ where χ is the indicator function, we see that ω has bounded variation as

$$\begin{aligned} V(\omega) &= \sum |\omega(x_{i+1} - x_i)| = \sum |\tilde{l}(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i))| \\ &= \sum s_i(\tilde{l}(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i)) = \tilde{l} \sum s_i(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i) \\ &\leq \|\tilde{l}\| \left\| \sum s_i(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i) \right\| = \|\tilde{l}\| \left\| \sum s_i \chi_{[x_i, x_{i+1}]}(x) \right\| \end{aligned}$$

where $s_i = \operatorname{sign}(\tilde{l}(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i))$. Then, for all $f \in \mathcal{C}([a,b])$ and a partition p, define

$$g_n = \sum f(x_i)(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i)).$$

In particular, we note that $\|g_n - f\|_\infty \to 0$ as $n \to \infty$ and thus,

$$l(f) = \tilde{l}(f) \leftarrow \tilde{l}(g_n) = \sum f(x_i)(\chi_{[a,t]}(x_{i+1}) - \chi_{[a,t]}(x_i)) = \int_a^b f \mathrm{d}\omega.$$

$\square$

4 Uniform Boundedness Principle

In this section we will prove the uniform boundedness principle, also known as the Banach-Steinhaus theorem. To achieve this we will first prove the Baire's category theorem. We will also provide some application of this principle.

4.1 Baire's Category Theorem

Definition 4.1 (Baire's Category)**.** Let (X, ρ) be a metric space. Then a subspace $M \subseteq X$ is said to be

- nowhere dense if $(\overline{M})^\circ = \emptyset$;
- of the 1^{st} category (or meager) if there exists a sequence of nowhere dense set $(M_n)_{n\in\mathbb{N}}$ such that $M = \bigcup_{n\in\mathbb{N}} M_n$;
- of the 2^{nd} category (or nonmeager) if it is not of the first category.

An example of a meager set is $\mathbb{Q} \subseteq \mathbb{R}$. Indeed, as $\mathbb{Q}$ is countable, it is a countable union of singletons, each of which are nowhere dense.

Theorem 8 (Baire's Category Theorem)**.** If a metric space $X \neq \emptyset$ is complete, then it is of the 2^{nd} category. Thus, if $X \neq \emptyset$ and

$$X = \bigcup_{k=1}^{\infty} A_k,$$

for some sequence of sets, then at least one A_k is not nowhere dense, i.e. contains a non-empty open subset.

Proof. Suppose otherwise, i.e. $X = \bigcup A_n$ where A_n is nowhere dense for all $n \in \mathbb{N}$. By definition $\overline{A_n}$ does not contain a non-empty open set and so, $\overline{A_n} \neq X$ (as X contains itself which is a non-empty open set) and $\overline{A_n}^c \neq \emptyset$ is open.

Let $p_1 \in \overline{A_1}^c$ and let $\epsilon_1 \in (0, 1/2)$ such that $B_{\epsilon_1}(p_1) \subseteq \overline{A_1}^c$. Now, as A_2 is nowhere dense, $B_{\epsilon_1}(p_1) \not\subseteq A_2$ and so $\overline{A_2}^c \cap B_{\epsilon_1}(p_1)$ is a non-empty open set. Then, we may choose $p_2 \in \overline{A_2}^c \cap B_{\epsilon_1}(p_1)$ and $\epsilon_2 \in (0, 1/4)$ such that $B_{\epsilon_2}(p_2) \subseteq \overline{A_2}^c \cap B_{\epsilon_1}(p_1)$. Repeating this process, we may find $p_{n+1} \in \overline{A_{n+1}}^c \cap B_{\epsilon_n}(p_n)$ and $\epsilon_{n+1} \in (0, 1/2^{n+1})$ such that

$$B_{\epsilon_{n+1}}(p_{n+1}) \subseteq \overline{A_{n+1}}^c \cap B_{\epsilon_n}(p_n).$$

It is clear that $(p_n)_{n\in\mathbb{N}}$ is Cauchy as for all $n, m \geq N$, $p_n, p_m \in B_{\epsilon_N}(p_N) \subseteq B_{1/2^N}(p_N)$ and thus,

$$d(p_n, p_m) \leq d(p_n, p_N) + d(p_N, p_m) < \frac{1}{2^N} + \frac{1}{2^N} = \frac{1}{2^{N-1}} \to 0$$

as $N \to \infty$. Hence, as X is complete, there exists some $p \in X$ such that $p_n \to p$ as $n \to \infty$. Now, by the triangle inequality, for all n,

$$d(x, x_n) \leq d(x, x_{n+k}) + d(x_{n+k}, x_n) < d(x, x_{n+k}) + \frac{1}{2^n},$$

and so, taking $k \to \infty$, $d(x, x_{n+k}) \to 0$ and so, $d(x, x_n) < 1/2^n$ implying $x \in B_{\epsilon_n}(p_n)$. But then, $x \in \bigcap B_{\epsilon_n}(p_n) \subseteq \bigcap \overline{A_n}^c$ implying $x \notin \bigcup A_n$, and hence, $\bigcup A_n \neq X$. $\square$

4.2 Banach-Steinhaus Theorem

Theorem 9 (Banach-Steinhaus)**.** Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of bounded linear operators from a Banach space X to a normed space Y such that for all $x \in X$, there exists $c_x \in (0, \infty)$ and

$$\sup_{n\in\mathbb{N}} \|T_n x\| \leq c_x.$$

Then, there exists some $c \in (0, \infty)$ such that

$$\sup_{n\in\mathbb{N}} \|T_n\| \leq c.$$

Proof. For $k \in \mathbb{N}$, define

$$A_k := \{x \in X \mid \|T_n x\| \leq k, n \in \mathbb{N}\}.$$

It is easy to see that A_k is closed for each k. Indeed, if $x_n \in A_k$ such that $x_n \to x \in X$, then $\|T_n x\| \leq k$ for all n by the continuity of the $\|T_n(\cdot)\|$. Furthermore, as $X = \bigcup A_k$ (as T_n is pointwise uniformly bounded), by the Baire's category theorem, there exists some $k_0 \in \mathbb{N}$, $x_0 \in X$, $r > 0$, such that

$$B_r(x_0) \subseteq A_{k_0}.$$

Then, for all $x \in X \setminus \{0\}$, let $\gamma_x := r/(2\|x\|)$ and $z := x_0 + \gamma x$. It is clear that $z \in B_r(x_0)$ and so, for all $n \in \mathbb{N}$,

$$k_0 \geq \|T_n z\| = \|T_n(x_0 + \gamma x)\| \geq |\gamma\|T_n x\| - \|T_n x_0\|| \geq r\frac{\|T_n x\|}{\|x\|} - \|T_n x_0\|.$$

Hence, $\|T_n x\|/\|x\| \leq (k_0 + \|T_n x_0\|)/r$, and since $\|T_n x_0\| \leq \sup_n \|T_n x_0\| \leq c_{x_0}$, we have the bound $\|T_n x\|/\|x\| \leq (k_0 + c_{x_0})/r$. Thus, n and x are arbitrary, we have

$$\sup_{n\in\mathbb{N}} \|T_n\| \leq \frac{1}{r}(k_0 + c_{x_0}).$$

$\square$

Proposition 4.1. Let X be the space of polynomials equipped with the norm $\|a_0 + a_1 t + \cdots + a_n t^n\| := \max_{i=0,\cdots,n} |a_i|$. Then X is not complete.

Proof. Define $T_n(p(t)) := \sum_{i=0}^n a_i$ where $p(t) = \sum_{i=0}^m a_i t^i$. Clearly, T_n is linear and for all $p(t) = \sum_{i=0}^m a_i t^i \in X$, $\|p\| = 1$, $a_i \leq 1$ for all $i = 1, \cdots, m$ implying $\|T_n(p)\| = |\sum^n a_i| \leq n+1$. Now, fixing $p(t) = \sum_{i=0}^m a_i t^i \in X$, we have $\|T_n p\| \leq \sum_{i=0}^m |a_i|$, and so, if X is complete, by the Banach-Steinhaus theorem, $\sup_n \|T_n\|$ is bounded. But for all n, we have

$$\|T_n(1 + x + \cdots + x^n)\| = \sum_{i=0}^n 1 = n + 1,$$

where $\|1 + x + \cdots + x^n\| = 1$, we have $\|T_n\| \geq n + 1$. Thus, $\sup_n \|T_n\| \geq \sup_n (n+1) = \infty$ implying $\sup_n \|T_n\|$ is not bounded and hence, X is not complete. $\square$

The Banach-Steinhaus theorem also allows us to show the existence of a continuous function which has Fourier series divergent at a point.

Let $u \in \mathcal{C}([0, 2\pi], \mathbb{R})$, and recall that it has Fourier coefficients defined by

$$a_m := \frac{1}{\pi} \int_0^{2\pi} u(t) \cos(mt) \mathrm{d}t; \ b_m := \frac{1}{\pi} \int_0^{2\pi} u(t) \sin(mt) \mathrm{d}t,$$

and its Fourier series is defined by

$$\mathcal{F}_u := \sum_{m \in \mathbb{N}} (a_m \cos(mt) + b_m \sin(mt)).$$

Let $X := \mathcal{C}([0, 2\pi], \mathbb{R})$ and define the sequence of linear operator T_n on X such that, for all $x \in X$,

$$T_n x := \frac{1}{2} a_0 + \sum_{m=1}^{n} a_m = \frac{1}{\pi} \int_0^{2\pi} x(t) \left(\frac{1}{2} + \sum_{m=1}^{n} \cos(mt) \right) \mathrm{d}t.$$

That is $T_n x$ is the Fourier series at $t = 0$ resumed only up to the n-th term. We will find some $x \in X$ such that $\|T_n x\|$ is not uniformly bounded, and so, the Fourier series of x diverges. Noting that

$$\begin{aligned} 2 \sin\left(\frac{1}{2} t\right) \sum_{m=1}^{n} \cos(mt) &= \sum_{m=1}^{n} \left(\sin\left(\left(m + \frac{1}{2}\right) t\right) - \sin\left(\left(m - \frac{1}{2}\right) t\right)\right) \\ &= \sin\left(\left(n + \frac{1}{2}\right) t\right) - \sin\left(\frac{1}{2}\right), \end{aligned}$$

as the second sum is telescoping, we have

$$1 + 2 \sum_{m=1}^{n} \cos(mt) = \frac{\sin\left(\left(n + \frac{1}{2}\right) t\right)}{\sin\left(\frac{1}{2} t\right)} =: q_n(t).$$

Thus,

$$T_n x = \frac{1}{2\pi} \int_0^{2\pi} x(t) q_n(t) \mathrm{d}t.$$

With this representation in mind, we see that

$$\|T_n x\| \leq \frac{1}{2\pi} \|x\| \int_0^{2\pi} |q_n(t)| \mathrm{d}t,$$

and so, $\|T_n\| \leq \frac{1}{2\pi} \int_0^{2\pi} |q_n(t)| \mathrm{d}t$. On the other hand, write $y(t) := \mathrm{sign}(q_n(t))$ so that $|q(t)| = |y(t) q(t)|$. While y is not continuous, it can be approximated by continuous functions of norm 1 in the sense that for all $\epsilon > 0$, there exists a continuous function x of norm 1 such that

$$\epsilon > \frac{1}{2\pi} \left| \int_0^{2\pi} (x(t) - y(t)) q_n(t) \mathrm{d}t \right| = \left| T_n x - \frac{1}{2\pi} \int_0^{2\pi} |q_n(t)| \right|.$$

Thus, as $\int_0^{2\pi} |q_n(t)| \mathrm{d}t \to \infty$ as $n \to \infty$, there must be some $x \in X$ such that $\|T_n x\|$ is not uniformly bounded as otherwise, this would contradict Banach-Steinhaus.

Let us now turn our focus back to ℓ_p spaces.

Proposition 4.2. Let $p, q \in (1, \infty)$ such that $1/p + 1/q = 1$ and suppose $(a_n)_{n\in\mathbb{N}}$ is a sequence of complex numbers such that $|\sum a_n x_n| < \infty$ for all $(x_n) \in \ell_p$. Then $(a_n) \in \ell_q$.

Proof. WLOG. let us assume (a_n) does not have any 0 terms. Let $f_n(x) := \sum_{i=0}^n a_i x_i$. It is not difficult to see that (f_n) are bounded linear and is pairwise uniformly bounded. Now, by Hölder, we have

$$|f_n(x)| \leq \left(\sum_{i=1}^n |a_i|^q\right)^{1/q} \|x\|p,$$

and so, $\|f_n\| \leq \left(\sum_{i=1}^n |a_i|^q\right)^{1/q}$. On the other hand, this bound is achieved by taking

$$x_i := \begin{cases} \frac{\overline{a_k}|a_i|^{q-2}}{(\sum_j^n |a_j|^q)^{1-q/q}}, & k = 1, \cdots, n, \\ 0, & \text{otherwise.} \end{cases}$$

So, $\left(\sum_{i=1}^n |a_i|^q\right)^{1/q} = \|f_n\|$ and by Banach-Steinhaus,

$$\|a\|_q = \left(\sum_{i=1}^\infty |a_i|^q\right)^{1/q} = \sup_n \|f_n\| < \infty.$$

□

4.3 Open Mapping Theorem

Definition 4.2. Let X, Y be metric spaces, then the function $T : D(T) \subseteq X \to Y$ is an open mapping if for all $U \subseteq D(T)$ open, $T(U)$ is open in Y.

Theorem 10 (The Open Mapping Theorem)**.** A bounded linear operator $T : X \to Y$ between Banach spaces X, Y is an open mapping if it is surjective.

Proof. Let $B_r^X(x), B_r^Y(y)$ denote open balls of radius r in X and Y respectively and B_r^X, B_r^Y denote open balls centred at the origin with radius r.

Since T is linear, it suffices to show that $T(B_1^X)$ contains an open ball containing 0 in Y. Indeed, for all open $A \subseteq D(T)$, if $y \in T(A)$, by definition, there exists some $x \in A$ such that $T(x) = y$. Now, as A is open, there exists some $B_r^X(x) \subseteq A$, and so $B_1^X = \frac{1}{r}(B_r^X(x) - x) \subseteq \frac{1}{r}(A - x)$. Then, by assumption, there exists some $\epsilon > 0$ such that

$$B_\epsilon^Y \subseteq T(B_1^X) \subseteq T\left(\frac{1}{r}(A - x)\right) = \frac{1}{r}(T(A) - T(x)) = \frac{1}{r}(T(A) - y).$$

Thus, $rB_\epsilon^Y + y \subseteq T(A)$. Now, since $y \in rB_\epsilon^Y + y \subseteq T(A)$ we have found an open ball containing y which is contained in $T(A)$. Hence, $T(A)$ is open and T is an open mapping.

We will first show $T(B_{1/2}^X(0))$ contains an open ball B^* not necessary centred at the origin. Consider

$$Y = T(X) = T\left(\bigcup_{k=1}^\infty kB_{1/2}(0)\right) = \bigcup_{k=1}^\infty kT(B_{1/2}(0)) = \bigcup_{k=1}^\infty \overline{kT(B_{1/2}(0))}.$$

By the Baire category theorem, at least on of the $\overline{kT(B_{1/2}(0))}$ must contain an open ball, and thus, by rescaling, we obtain an open ball $B^* = B_\epsilon(y_0) \subseteq \overline{T(B_{1/2}(0))}$.

Now, as $y_0 \in B_\epsilon(y_0) \subseteq \overline{T(B_{1/2}(0))}$ and, for all $y \in B_\epsilon(0) = B_\epsilon(y_0) - y_0$, we have $y + y_0 \in B_\epsilon(y_0) \subseteq \overline{T(B_{1/2}(0))}$, taking sequences $(u_n), (v_n) \subseteq T(B_{1/2}(0))$ such that $u_n \to y + y_0$ and $v_n \to y_0$, there exists some $(w_n), (z_n) \subseteq B_{1/2}(0)$ such that $Tw_n = u_n \to y + y_0, Tz_n = v_n \to y_0$. Then, by sequential continuity

$$T(w_n - z_n) = Tw_n - Tz_n = u_n - v_n \to y,$$

and $w_n - z_n \in B_1(0)$ as $\|w_n - z_n\| \leq \|w_n\| + \|z_n\| < 1/2 + 1/2 = 1$. Hence, $y \in \overline{T(B_1(0))}$ and thus, $B_\epsilon(0) \subseteq \overline{T(B_1(0))}$.

Finally, we will show that $B_{\epsilon/2}(0) \subseteq T(B_1(0))$. By linearity, we have

$$V_n := B_{\epsilon/2^n}(0) \subseteq 2^{-n}\overline{T(B_1(0))} = \overline{T(B_{1/2^n}(0))} =: \overline{T(B_n)}.$$

Let $y \in B_{\epsilon/2} = V_1$, then $y \in \overline{T(B_1)}$ and thus, there exists some $x_1 \in B_1$ such that $\|y - Tx_1\| < \epsilon/4$. This means that $y - Tx_1 \in V_2 \subseteq \overline{T(B_2)}$ and there exists some $x_2 \in B_2$, $\|y - Tx_1 - Tx_2\| < \epsilon/8$. Repeating this process, we find a sequence (x_n) such that $x_n \in B_n$ and

$$\left\| y - \sum_{k=1}^n Tx_k \right\| < \frac{\epsilon}{2^{n+1}}.$$

Considering for $n > m$,

$$\left\| \sum_{k=1}^n Tx_k - \sum_{k=1}^m Tx_k \right\| = \left\| \sum_{k=m+1}^n Tx_k \right\| \leq \sum_{k=m+1}^n \|Tx_k\| \leq \sum_{k=m+1}^\infty \|Tx_k\| < \sum_{k=m+1}^\infty \frac{1}{2^k} \to 0,$$

as $m \to \infty$, we have $S_n := \sum_{k=1}^n Tx_k$ is Cauchy, and thus converges to some $s \in Y$. As $\|y - S_n\| < \epsilon/2^{n+1}$, $\|y - s\| \leq \|y - S_n\| + \|S_n - s\| \to 0$ as $n \to \infty$, $\|y - s\| = 0$ and thus, $y = s$.

Now, $\sum_{k=1}^n x_k$ converges to some $x \in X$ as it is absolutely convergent since $x_n \in B_n$ implying $\|x_n\| < 1/2^n$ and so $\sum \|x_n\| < \sum 1/2^n = 1 < \infty$. Thus, by sequential continuity, we have $T(\sum_{k=1}^n x_k) \to Tx$ as $n \to \infty$ and so, $Tx = y$. Finally, $x \in B_1(0)$ since addition is continuous, and so

$$\|x\| = \lim_{n\to\infty} \left\| \sum_{k=1}^n x_n \right\| \leq \lim_{n\to\infty} \sum_{k=1}^n \|x_n\| < \lim_{n\to\infty} \sum_{k=1}^n \frac{1}{2^k} = 1,$$

implying $x \in B_1(0)$ and thus, as $y = Tx$, $y \in T(B_1(0))$. $\square$

With the open mapping theorem, we see that a bijective bounded linear operator is automatically a homeomorphism and T^{-1} is bounded.

4.4 Closed Graph Theorem

Definition 4.3 (Closed Linear Operator)**.** Let X, Y be normed spaces and $T : D(T) \subseteq X \to Y$ be a linear operator. The graph of T is defined to be the set

$$\mathcal{G}(T) := \{(x, Tx) \in X \times Y \mid x \in D(T)\}.$$

T is said to be a closed linear operator if its graph is closed in the normed space equipped with the norm $\|(x, y)\| = \|x\|_X + \|y\|_Y$.

Proposition 4.3. Let X, Y be normed spaces and $T : D(T) \subseteq X \to Y$ be a linear operator. Then T is a closed linear operator if and only if for all $(x_n) \subseteq D(T)$, $x_n \to x$ and $Tx_n \to y$, we have $x \in D(T)$ and $Tx = y$.

Proof. Suppose $(x_n, Tx_n) \in \mathcal{G}(T)$ such that $(x_n, T_n) \to (x, y) \in X \times Y$. Then

$$0 \leftarrow \|(x_n, Tx_n) - (x, y)\| = \|(x_n - x, Tx_n - y)\| = \|x_n - x\|_X + \|Tx_n - y\|_Y,$$

and so, $x_n \to x$ and $Tx_n \to y$ and thus, $Tx = y$ implying $(x, y) \in \mathcal{G}(T)$ and $\mathcal{G}(T)$ is closed.

Conversely, if $\mathcal{G}(T)$ is closed, then for all $x_n \to x$, $Tx_n \to y$, (x, y) is a limit point of $\mathcal{G}(T)$. Thus,, as $\mathcal{G}(T)$ is closed, $(x, y) \in \mathcal{G}(T)$ and so, $y = Tx$. $\square$

Corollary 4.1. Let X, Y be normed spaces and $T : D(T) \subseteq X \to Y$ be a continuous linear operator. Then T is a closed linear operator if $D(T)$ is closed.

Proof. Suppose $x_n \to x \in X$ and $Tx_n \to y$. As $D(T)$ is closed, $x \in D(T)$, and by sequential continuity, $Tx = y$. $\square$

Theorem 11 (Closed Graph Theorem)**.** Let X, Y be Banach spaces and $T : D(T) \subseteq X \to Y$ be a closed linear operator. Then if $D(T)$ is closed in X, then T is bounded.

Proof. We first observe that $X \times Y$ is complete. Indeed, if $((x_n, y_n))_{n \in \mathbb{N}}$ is Cauchy, then so are $(x_n), (y_n)$. Thus, as both X, Y are Banach, there exists some $x \in X, y \in Y$, $x_n \to x, y_n \to y$ and so,

$$\|(x_n, y_n) - (x, y)\| = \|x_n - x\| + \|y_n - y\| \to 0.$$

Furthermore, since $\mathcal{G}(T)$ and $D(T)$ are by assumption closed, they are also complete as they are closed subspaces of a Banach space.

Now, define

$$P : \mathcal{G}(T) \to D(T) : (x, Tx) \to x.$$

It is clear that P is linear and is bounded since for all $(x, Tx) \in \mathcal{G}(T)$, $\|(x, Tx)\| = 1$, we have $1 = \|(x, Tx)\| = \|x\| + \|Tx\| \geq \|x\| = \|P(x, Tx)\|$. Furthermore, P is bijective with the inverse $P^{-1}(x) = (x, Tx)$ which is also a bounded linear operator (as the inverse of a bounded linear operator is bounded). But, this implies

$$\|Tx\| \leq \|x\| + \|Tx\| = \|(x, Tx)\| = \|P^{-1}(x)\| \leq \|P^{-1}\|\|x\|,$$

proving T is bounded. $\square$

5 Spectral Theory

5.1 Compact Operators

Definition 5.1. Let X, Y be normed spaces. An operator $T : X \to Y$ is a compact operator if for all $B \subseteq X$ bounded, $\overline{T(B)}$ is compact.

Proposition 5.1. Let X, Y be normed spaces and $T : X \to Y$ be an operator. Then T is compact if and only if for all bounded sequence (x_n) of X, (Tx_n) has a convergent subsequence.

Proof. Follows by considering the criterion for relatively compactness in metric spaces. In particular, given a subset B of a metric space, $\overline{B}$ is compact if and only if all sequences in B contain a subsequence which converges in $\overline{B}$. Indeed, if (x_n) is a sequence in $\overline{B}$, then we may construct a sequence $(y_n) \subseteq B$ such that $d(x_n, y_n) < 1/n$. Then, if $(y_{n_i}) \subseteq (y_n)$ such that $y_{n_i} \to y \in \overline{B}$,

$$d(x_{n_i}, y) \le d(x_{n_i}, y_{n_i}) + d(y_{n_i}, y) < \frac{1}{n_i} + d(y_{n_i}, y) \to 0$$

as $i \to \infty$. Thus, $\overline{B}$ is compact. □

Proposition 5.2. Let X, Y be normed spaces. Then the set of compact operators from X to Y form an linear space.

Proof. Exercise. □

Proposition 5.3. A linear operator $T : X \to Y$ where X, Y are normed spaces is compact if T is bounded and $\dim T(X) < \infty$ (such operators are known as finite rank operators). Thus, if X is finite-dimensional, every bounded linear operator is compact.

Proof. Recall that for finite dimensional spaces, closed and bounded implies compactness. Thus, for all bounded B, as $T(B)$ is bounded, $\overline{T(B)}$ is closed and bounded and hence, T is compact. □

Proposition 5.4. Let (T_n) be a sequence of compact linear operators from a normed space X to a Banach space Y. Then if $T_n \to T$ with respect to the operator norm, T is also compact.

Proof. Let (x_n) be a bounded sequence in X by some $c \in \mathbb{R}^+$. By assumption, as T_1 is compact, there exists a subsequence $(x_n^{(1)})$ of (x_n) such that $T_1 x_n^{(1)}$ converges. Then, as T_2 is compact, there is a subsequence $(x_n^{(2)})$ of $(x_n^{(1)})$ such that $(T_2 x_n^{(2)})$ converges as well. Repeating this process, we can find $(x_n^{(k)})$ a subsequence of $(x_n^{(l)})$ for all $l \le k$ such that $T_i x_n^{(k)}$ converges for $i = 1, 2, \cdots, k$. Now, defining $y_n := x_n^n$, I claim (y_n) is a subsequence of (x_n) such that $T_i y_n$ converges for all $i \in \mathbb{N}$. Indeed, for all $i \in \mathbb{N}$, once $n \ge i$, the remaining sequence of (y_n) is contained $(x_n^{(i)})$. Thus, $T_i y_n$ is eventually contained in $T_i x_n^{(i)}$ which converges.

Now, as $T_n \to T$, for all $\epsilon > 0$, there exists N such that for all $n \geq N$, $\|T_n - T\| < \epsilon/(3c)$. On the other hand, as $(T_N y_n)$ is convergent, it is Cauchy, and so, there exists sufficiently large M, such that $\|T_N y_i - T_N y_j\| < \epsilon/3$ for all $i, j \geq M$. Hence, for all $i, j \geq M$,

$$\|Ty_i - Ty_j\| \leq \|Ty_i - T_N y_i\| + \|T_N y_i - T_N y_j\| + \|T_N y_j - Ty_j\| < \epsilon.$$

Therefore (Ty_n) is Cauchy and so, T is compact. □

5.1.1 Hilbert-Schmidt Integral Operator

Definition 5.2 (Hilbert-Schmidt Integral Operator)**.** The Hilbert-Schmidt integral operator $T : H \to H$ where $H := L^2(\mathbb{R}^d)$ is the linear operator

$$Tf : x \mapsto \int_{\mathbb{R}^d} K(x, y) f(y) \mathrm{d}\lambda(y),$$

for some kernel map $K \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

The map $y \mapsto K(x, y)f(y)$ is integrable for all x since by Fubini, the map $y \mapsto K(x, y)$ is square integrable for all x and thus,

$$\|K(x, \cdot)f(\cdot)\|_1 = \langle K(x, \cdot), f\rangle^2 \leq \|K(x, \cdot)\|_2 \|f\|_2.$$

We note that this also implies the operator is bounded with the bound $\|T\| \leq \|K\|_2$.