Abstract Functional Analysis

by, Warren B. Moors

Department of Mathematics
The University of Auckland
Auckland
New Zealand

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

Contents

Chapter 1

Zorn's Lemma

A **partially ordered set** $(X, \leqslant)$ is a set X with a binary relation "$\leqslant$" satisfying the following three axioms:

(i) for every $x \in X$, $x \leqslant x$;

(ii) for every $x, y \in X$, if $x \leqslant y$ and $y \leqslant x$, then $x = y$;

(iii) for every $x, y, z \in X$, if $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$.

An element x of a partially ordered set $(X, \leqslant)$ is called **maximal** if there are no other elements greater than it, i.e., if x is maximal, then for every $y \in X$, if $x \leqslant y$, then $x = y$.

Example 1.1. *Let Y be a nonempty set and let X be the set of all nonempty proper subsets of Y. Define "$\leqslant$" on X by, $A \leqslant B$ if, and only if, $A \subseteq B$. Then $(X, \leqslant)$ is a partially ordered set.*

Exercise 1.2. *Find the maximal elements in the partially ordered set $(X, \leqslant)$ described above.*

A **totally ordered set** $(X, \leqslant)$ is a set X with a binary relation "$\leqslant$" satisfying the following three axioms:

(i) for every $x, y \in X$, either $x \leqslant y$, or $y \leqslant x$;

(ii) for every $x, y \in X$, if $x \leqslant y$ and $y \leqslant x$, then $x = y$;

(iii) for every $x, y, z \in X$, if $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$.

Example 1.3. *If $(X, \leqslant)$ is a totally ordered set, then $(X^2, \preceq)$ is also a totally ordered set if "$\preceq$" is defined by, $(x, y) \preceq (x', y')$ if, and only if, $x < x'$ or $x = x'$ and $y \leqslant y'$.*

Exercise 1.4. *Show that if $(T, \leqslant)$ is a totally ordered set and T has only finitely many elements, then T has a* **largest element** *i.e., there exists an element $t_{\max} \in T$ such that $t \leqslant t_{\max}$ for all $t \in T$.*

We will say that a subset S of a partially ordered set $(X, \leqslant)$ is **bounded above** if there exists an element $x \in X$ such that $s \leqslant x$ for all $s \in S$.

Theorem 1.5 (Zorn's Lemma)**.** *Let $(X, \leqslant)$ be a nonempty partially ordered set. If every totally ordered subset of X is bounded above, then $(X, \leqslant)$ has a maximal element.*

Remarks 1.6. *Zorn's Lemma is equivalent to the "Axiom of Choice".*

Exercise 1.7. *Let I be a proper ideal in a commutative ring with identity $\langle R, +, \cdot \rangle$. Show that I is contained in a maximal proper ideal in R, i.e., show that every proper ideal is contained in a maximal proper ideal.*

Vector spaces

A **vector space** $(V; +; \cdot)$ over a field $\mathbb{K}$ is a set V together with two binary operations $+ : V \times V \to V$ and $\cdot : \mathbb{K} \times V \to V$ which obey the following set of rules:

1. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$;
2. $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$;
3. there exists an element $\boldsymbol{O} \in V$ such that $\boldsymbol{u} + \boldsymbol{O} = \boldsymbol{O} + \boldsymbol{u} = \boldsymbol{u}$ for all $\boldsymbol{u} \in V$;
4. for each $\boldsymbol{u} \in V$ there exists an element $\boldsymbol{v} \in V$ such that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u} = \boldsymbol{O}$;
5. $t \cdot (\boldsymbol{u} + \boldsymbol{v}) = t \cdot \boldsymbol{u} + t \cdot \boldsymbol{v}$ for each $t \in \mathbb{K}$ and all elements $\boldsymbol{u}, \boldsymbol{v} \in V$;
6. $(s + t) \cdot \boldsymbol{u} = s \cdot \boldsymbol{u} + t \cdot \boldsymbol{u}$ for each $\boldsymbol{u} \in V$ and all s and $t \in \mathbb{K}$;
7. $(st) \cdot \boldsymbol{u} = s \cdot (t \cdot \boldsymbol{u})$ for each $\boldsymbol{u} \in V$ and all s and $t \in \mathbb{K}$;
8. $1 \cdot \boldsymbol{u} = \boldsymbol{u}$ for each $\boldsymbol{u} \in V$.

The elements of the set V are called **vectors** and the operations $+$ and $\cdot$ are called **vector addition** and **scalar multiplication** respectively. The vector $\boldsymbol{O}$ is called the **zero vector**.

Example 1. The set of all geometric vectors in 2-space (or 3-space) with the operations of vector addition and scalar multiplication, as defined in first year.

Example 2. The collection of all ordered n-tuples of elements of $\mathbb{K}$, together with the operations of component-wise addition and scalar multiplication, i.e.,

$$\begin{aligned}(a_1, a_2, \ldots a_n) + (b_1, b_2, \ldots b_n) &:= (a_1 + b_1, a_2 + b_2, \ldots a_n + b_n) \\ t \cdot (a_1, a_2, \ldots a_n) &:= (ta_1, ta_2, \ldots ta_n)\end{aligned}$$

We shall denote this system by $\mathbb{K}^n$.

Example 3. Let X be a nonempty set. Then the system $(F(X); +; \cdot)$ comprised of all the $\mathbb{K}$-valued functions defined on X (i.e., $F(X)$), together with the operations of pointwise addition and pointwise scalar multiplication, i.e., if $f, g \in F(X)$ then $f + g \in F(X)$ is defined by, $(f + g)(x) := f(x) + g(x)$ for each $x \in X$ and if $t \in \mathbb{K}$ then $t \cdot f \in F(X)$ is defined by, $(t \cdot f)(x) := t \cdot f(x)$ for each $x \in X$.

Example 4. Let X be a nonempty set. Then the system $(F_0(X); +; \cdot)$ comprised of all the $\mathbb{K}$-valued functions defined on X with finite support (i.e., if $f \in F_0(X)$ then $f \in F(X)$

and $\{x \in X : f(x) \neq 0\}$ is a finite set), together with the operations of pointwise addition and pointwise scalar multiplication (as in Example 3.).

Given two vector spaces $(V'; \oplus; \odot)$ and $(V; +; \cdot)$ we say that $(V'; \oplus; \odot)$ is **isomorphic to** $(V; +; \cdot)$ if there exists a 1-to-1 and onto mapping $\varphi : V' \to V$ such that (i) $\varphi(\boldsymbol{u} \oplus \boldsymbol{v}) = \varphi(\boldsymbol{u}) + \varphi(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v} \in V'$ and (ii) $\varphi(t \odot \boldsymbol{u}) = t \cdot \varphi(\boldsymbol{u})$ for all $t \in \mathbb{K}$ and all $\boldsymbol{u} \in V'$.

Example 1. The geometric vectors in 2-space are isomorphic to $\mathbb{R}^2$. To see this, let $\mathcal{S}$ be a basis for 2-space. Then the mapping φ that maps each vector $\boldsymbol{u}$ in 2-space to its $\mathcal{S}$-coordinates fulfils the hypotheses above.

Example 2. The geometric vectors in 3-space are isomorphic to $\mathbb{R}^3$. To see this, let $\mathcal{S}$ be a basis for 3-space. Then the mapping φ that maps each vector $\boldsymbol{u}$ in 3-space to its $\mathcal{S}$-coordinates fulfils the hypotheses above.

Example 3. Every vector space $(V; +; \cdot)$ over the real numbers, that consists of more than just the zero vector, is isomorphic to $(F_0(X); +; \cdot)$ for some nonempty set X.

A **linear combination** of elements $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ of a vector space V with coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, is an expression of the form: $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \ldots + \lambda_n \boldsymbol{x}_n$ (or rather, the value of this expression).

Exercise 1.8. *Show that if* $(V; +; \cdot)$ *is a vector space and* $\mathscr{F}$ *is a family of subspaces of* V, *then* $\bigcap_{S \in \mathscr{F}} S$ *is a subspace of* $(V; +; \cdot)$.

The **span** of a subset $X \subseteq V$, denoted **span**(X), is the smallest subspace of V containing the set X. This is,

$$\operatorname{span}(X) = \bigcap \{S \in 2^V : X \subseteq S \text{ and } S \text{ is a subspace of } V\}.$$

In particular, $\operatorname{span}(\varnothing) = \{\boldsymbol{O}\}$.

Exercise 1.9. *Let* X *be a nonempty subset of a vector space* V. *Show that* $\operatorname{span}(X)$ *is the set of all elements of* V *that can be expressed as a linear combination of elements of* X.

A nonempty finite subset $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ of V is said to be **linearly independent** if the only solution to the equation $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \ldots + \lambda_n \boldsymbol{x}_n = \mathbf{0}$ is $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Otherwise, the set $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ is said to be **linearly dependent**. An arbitrary subset $X \subseteq V$ is said to be **linearly independent** if every nonempty finite subset of X is linearly independent. So vacuously, $\varnothing$ is linearly independent. A subset $X \subseteq V$ is termed a **basis** for V if it is linearly independent and **spans** V, i.e., $\operatorname{span}(X) = V$.

Basic facts about bases

(i) every element $\boldsymbol{x} \in V$ admits a unique **basis decomposition**, this is, every $\boldsymbol{x} \in V$ can be *uniquely* expressed as a linear combination of elements of a fixed basis X;

(ii) if Y spans V, then Y contains a basis for V;

(iii) in particular, every nonzero vector space admits a basis;

(iv) every linearly independent subset Y can be extended to form a basis for V.

A vector space V is called **finite dimensional** if it admits a basis with only finite many elements. If a vector space is not finite dimensional, then it is called **infinite dimensional**.

A function from one vector space to another is called an **operator** (or **transformation**). A mapping from a vector space (over a field $\mathbb{K}$) into the field $\mathbb{K}$ is called a **functional**. An operator $f : U \to V$ is called a **linear** operator if for any $\boldsymbol{x}, \boldsymbol{y} \in U$ and $\lambda \in \mathbb{K}$, $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ and $f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$. The collection of all linear functionals on a vector space V forms a subspace of the vector space $\mathbb{K}^V$, under pointwise addition and pointwise scalar multiplication. It is denoted $V^{\#}$ and is called the **algebraic dual** of V. If V is finite dimensional, then V is isomorphic to $V^{\#}$.

Theorem 1.10. *Every nonzero vector space* $(V; +; \cdot\,)$ *admits a basis.*

Proof. Let $(V; +; \cdot\,)$ be a nonzero vector space and let X be the family of all linearly independent subsets of V. Then $X \neq \varnothing$ and $(X, \subseteq)$ is a partially ordered set (Note: if $\boldsymbol{x} \in V \backslash \{\mathbf{0}\}$, then $\{\boldsymbol{x}\} \in X$). We claim that X contains a maximal element. By Zorn's Lemma to show this we need only show that each totally ordered subset of X has an upper bound. Let $\varnothing \neq T \subseteq X$ be totally ordered and let $U := \bigcup\{I : I \in T\}$. Clearly $I \subseteq U$ for each $I \in T$ and so U is an upper bound for T, provided we have $U \in X$. So suppose $\boldsymbol{x}_j \in U$, $1 \leqslant j \leqslant n$. Then for each $1 \leqslant j \leqslant n$ there exists a $I_j \in T$ such that $\boldsymbol{x}_j \in I_j$. Now since T is totally ordered their exists a $k \in \{1, 2, \dots, n\}$ so that $I_j \subseteq I_k$ for each $1 \leqslant j \leqslant n$. Hence $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n\} \subseteq I_k$ and so are linearly independent. This shows that $U \in X$. Let $X_{\max}$ be a maximal element in $(X, \subseteq)$. We claim that $\mathrm{span}(X_{\max}) = V$, for if this is not the case, then we may take $\boldsymbol{x} \in V \backslash \mathrm{span}(X_{\max})$ and set $X^* := X_{\max} \cup \{\boldsymbol{x}\}$. Then $X^* \in X$, $X_{\max} \subseteq X^*$ but $X_{\max} \neq X^*$; which contradicts the maximality of $X_{\max}$. Hence, $X_{\max}$ is a basis for V. $\square$

Note that if $V = \{\boldsymbol{O}\}$, then technically $\varnothing$ is a basis for V as $\varnothing$ is linearly independent and $\mathrm{span}(\varnothing) = \{\boldsymbol{O}\} = V$.

Exercise 1.11. *Prove that every vector space* $(V; +; \cdot\,)$ *over the real numbers, that consists of more than just the zero vector, is isomorphic to* $(F_0(X); +; \cdot\,)$ *for some nonempty set* X*. This is the first "Representation Theorem" contained in this course.*

Exercise 1.12. *Prove that every linearly independent subset* Y *of a nonzero vector space* $(V; +; \cdot\,)$ *can be extended to form a basis for* V*.*

Exercise 1.13. *Prove that if* Y *spans a nonzero vector space* $(V; +; \cdot\,)$*, then* Y *contains a basis for* V*.*

Chapter 2

Introduction to Banach spaces

A **norm** on a vector space V (over a field $\mathbb{K}$) is a function, denoted by $\|\cdot\|$, from V into $\mathbb{R}$ such that:

(i) $\|x\| \geqslant 0$ for all $x \in V$ and $\|x\| = 0$ if, and only if, $x = 0$;

(ii) $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in V$ and all $\lambda \in \mathbb{K}$;

(iii) $\|x + y\| \leqslant \|x\| + \|y\|$ for all $x, y \in V$.

Any pair $(X, \|\cdot\|)$ consisting of a vector space and a norm is called a **normed linear space**.

Proposition 2.1. *Let $(X, \|\cdot\|)$ be a normed linear space. Then the function $\rho : X^2 \to [0, \infty)$ defined by, $\rho(x, y) := \|x - y\|$ for all $x, y \in X$ defines a metric on X.*

Proof. From the definition, $\rho(x, y) = 0$ if, and only if, $\|x - y\| = 0$ and this only occurs when $x = y$. Again, directly from the definition, if $x, y \in X$, then

$$\rho(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1|\|y - x\| = \|y - x\| = \rho(y, x).$$

So it remains to verify the triangle inequality. Let x, y and z be members of X, then

$$\rho(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leqslant \|x - y\| + \|y - z\| = \rho(x, y) + \rho(y, z).$$

This completes the proof. $\square$

In a normed linear space $(X, \|\cdot\|)$ we shall denote by, $B_X := \{x \in X : \|x\| \leqslant 1\}$ and $S_X := \{x \in X : \|x\| = 1\}$. For a subset A of a vector space V (over a field $\mathbb{K}$) and a scalar $\lambda \in \mathbb{K}$ we define $\lambda A := \{x \in V : x = \lambda a \text{ for some } a \in A\}$. If $x_0 \in V$, then we define $x_0 + A := \{x \in V : x = x_0 + a \text{ for some } a \in A\}$.

Proposition 2.2. *Let $(X, \|\cdot\|)$ be a normed linear space. Then for each $x \in X$ and each positive real number r, $x + rB_X = B[x; r] := \{y \in X : \|y - x\| \leqslant r\}$.*

Proof. Suppose that $y \in x + rB_X$, then $(y - x) \in rB_X$ and so $(1/r)(y - x) \in B_X$; which implies that $\|(1/r)(y - x)\| = |1/r|\|x - y\| \leqslant 1$, i.e., $\|y - x\| \leqslant r$. Therefore, $\rho(x, y) \leqslant r$ and so $y \in B[x; r]$. Conversely, suppose that $y \in B[x; r]$, then $\|y - x\| \leqslant r$ and so $\|(1/r)(y-x)\| \leqslant 1$, i.e., $(1/r)(y-x) \in B_X$. Therefore, $(y-x) \in rB_X$ and so $y \in x+rB_X$. This shows that $B[x; r] = x + rB_X$. $\square$

A **Banach space** $(X, \|\cdot\|)$ is a normed linear space that is complete in the metric defined by, $\rho(x, y) := \|x - y\|$, (i.e., Cauchy sequences in (X, ρ) are convergent).

Let $(X, \|\cdot\|)$ be a normed linear space. We say that a **series** $\sum_{k=1}^{\infty} x_k$ in X (i.e., $x_k \in X$ for all $k \in \mathbb{N}$) is **convergent** if the sequence (of partial sums) $s_n := \sum_{k=1}^{n} x_k$ is convergent in X. We say that a series $\sum_{k=1}^{\infty} x_k$ is **absolutely convergent** if $\sum_{k=1}^{\infty} \|x_k\|$ is convergent.

Proposition 2.3. *A normed linear space $(X, \|\cdot\|)$ is a Banach space if, and only if, every absolutely convergent series in $(X, \|\cdot\|)$ is convergent.*

Proof. Suppose that $(X, \|\cdot\|)$ is a Banach space and $\sum_{k=1}^{\infty} x_k$ is an absolutely convergent series in $(X, \|\cdot\|)$. For each $n \in \mathbb{N}$, let $s_n := \sum_{k=1}^{n} x_k$ and $t_n := \sum_{k=1}^{n} \|x_k\|$. Then, for any $(m, n) \in \mathbb{N}^2$ with $m < n$ we have that

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leqslant \sum_{k=m+1}^{n} \|x_k\| = |t_n - t_m|.$$

Since the sequence $(t_n : n \in \mathbb{N})$ is convergent it is also Cauchy. It then follows that the sequence $(s_n : n \in \mathbb{N})$ is a Cauchy sequence in $(X, \|\cdot\|)$ and hence convergent.

Converse: Suppose that $(X, \|\cdot\|)$ is a normed linear space in which every absolutely convergent series in $(X, \|\cdot\|)$ is convergent. Let $(x_n : n \in \mathbb{N})$ be a Cauchy sequence in $(X, \|\cdot\|)$. To show that $(x_n : n \in \mathbb{N})$ is convergent it is sufficient to show that it possesses a convergent subsequence. To this end, let us inductively define a strictly increasing sequence $(n_k : k \in \mathbb{N})$ of natural numbers such that $\sup\{\|x_i - x_j\| : n_k \leqslant i, j \in \mathbb{N}\} < 1/k^2$. Then define, $(y_k : k \in \mathbb{N})$ in X by, $y_k := x_{n_{k+1}} - x_{n_k}$. By construction the series $\sum_{j=1}^{\infty} y_j$ is absolutely convergent, and hence by assumption, convergent. Let us also note that $x_{n_1} + \sum_{j=1}^{k} y_j = x_{n_{k+1}}$ for all $k \in \mathbb{N}$. Therefore, $(x_{n_k} : k \in \mathbb{N})$ is a convergent subsequence of $(x_n : n \in \mathbb{N})$; which completes the proof. $\square$

Theorem 2.4. *Let $(X, \|\cdot\|)$ be a Banach space and let Y be a subspace of $(X, \|\cdot\|)$. Then $(Y, \|\cdot\|)$ is a Banach space if, and only if, Y is a closed subspace of $(X, \|\cdot\|)$.*

Proof. The proof that a closed subspace of a Banach space is again a Banach space is left as an easy exercise for the reader. To prove the converse it suffices to show that $\overline{Y} \subseteq Y$. So let $y \in \overline{Y}$. Then there exists a sequence $(y_n : n \in \mathbb{N})$ in Y converging to y. Therefore, $(y_n : n \in \mathbb{N})$ is a Cauchy sequence in $(Y, \|\cdot\|)$. Now since $(Y, \|\cdot\|)$ is a Banach space there exists a point $y_\infty \in Y$ such that $\lim_{n\to\infty} y_n = y_\infty$ (the limit is considered in $(Y, \|\cdot\|)$). On the other hand, $\lim_{n\to\infty} y_n = y_\infty$ (considered in $(X, \|\cdot\|)$). Since the limit of a convergent sequence in $(X, \|\cdot\|)$ is unique, $y = y_\infty \in Y$. Hence, $\overline{Y} \subseteq Y$. $\square$

Let Y be a closed subspace of a normed linear space $(X, \|\cdot\|)$. For each $x \in X$ we consider the coset $\widehat{x}$ relative to Y, $\widehat{x} := x + Y$. The space $X/Y := \{\widehat{x} : x \in X\}$ of all cosets, together with the addition and scalar multiplication defined by, $\widehat{x} + \widehat{y} = \widehat{x + y}$ and $\lambda\widehat{x} = \widehat{\lambda x}$ is a vector space. It is routine to check that $\|\widehat{x}\| := \inf\{\|y\| : y \in \widehat{x}\}$ defines a norm on X/Y.

Let Y be a closed subspace of a normed linear space $(X, \|\cdot\|)$. Then the space X/Y endowed with the norm $\|\widehat{x}\| = \inf\{\|y\| : y \in \widehat{x}\}$ is called the **quotient space of X with respect to Y**.

Exercise 2.5. *Let Y be a closed subspace of a normed linear space $(X, \|\cdot\|)$. Show that the mapping $x \mapsto \widehat{x}$ from $(X, \|\cdot\|)$ into $(X/Y, \|\cdot\|)$ is linear and continuous.*

Theorem 2.6. *Let Y be a closed subspace of a Banach space $(X, \|\cdot\|)$. Then $(X/Y, \|\cdot\|)$ is a Banach space.*

Proof. Let $\sum_{k=1}^{\infty} \widehat{x}_k$ be an absolutely convergent series in X/Y. For each $k \in \mathbb{N}$, choose $y_k \in \widehat{x}_k$ so that $\|\widehat{x}_k\| \leqslant \|y_k\| < \|\widehat{x}_k\| + 1/k^2$. Then $\sum_{k=1}^{\infty} \|y_k\|$ is convergent. Since $(X, \|\cdot\|)$ is a Banach space $\sum_{k=1}^{\infty} y_k$ is convergent in X. Let $y := \sum_{k=1}^{\infty} y_k$, then

$$\widehat{y} = \widehat{\lim_{n\to\infty} \sum_{k=1}^{n} y_k} = \lim_{n\to\infty} \widehat{\sum_{k=1}^{n} y_k} = \lim_{n\to\infty} \sum_{k=1}^{n} \widehat{y_k} = \lim_{n\to\infty} \sum_{k=1}^{n} \widehat{x}_k = \sum_{k=1}^{\infty} \widehat{x}_k.$$

This shows that every absolutely convergent series in $(X/Y, \|\cdot\|)$ is convergent; thus $(X/Y, \|\cdot\|)$ is a Banach space. $\square$

Next, we examine finite dimensional normed linear spaces.

Let $\|\cdot\|$ and $|||\cdot|||$ be norms on a vector space V. We say that the norm $\|\cdot\|$ is **equivalent** to the norm $|||\cdot|||$ if, and only if, there exists real numbers $0 < m \leqslant M < \infty$ such that $m\|x\| \leqslant |||x||| \leqslant M\|x\|$ for all $x \in V$.

Exercise 2.7. *Let $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_3$ be norms on a vector space V. Show that if $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$ and $\|\cdot\|_2$ is equivalent to $\|\cdot\|_3$, then $\|\cdot\|_1$ is equivalent to $\|\cdot\|_3$. Also show that $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$ if, and only if, $\|\cdot\|_2$ is equivalent to $\|\cdot\|_1$.*

Theorem 2.8 (Fundamental Theorem of Finite Dimensional Normed Linear Spaces)**.** *Let $\|\cdot\|$ and $|||\cdot|||$ be norms on a finite dimensional vector space V. Then $\|\cdot\|$ and $|||\cdot|||$ are equivalent norms (i.e., all norms on a finite dimensional space are equivalent).*

Proof : Let $\mathscr{B} := \{e_1, e_2, \ldots, e_n\}$ be a basis for V. On V we define the $\|\cdot\|_1$ norm by, $\|x\|_1 := \sum_{k=1}^{n} |x_k|$ where $(x_k : 1 \leqslant k \leqslant n)$ are the coordinates of x with respect to $\mathscr{B}$. It is easy to show that $\|\cdot\|_1$ is indeed a norm on V. So it will be sufficient to show that if $\|\cdot\|$ is any norm on V, then $\|\cdot\|$ is equivalent to $\|\cdot\|_1$. Let $M := \max\{\|e_k\| : 1 \leqslant k \leqslant n\}$.

Then for any $x \in V$,

$$\begin{aligned}
\|x\| &= \left\|\sum_{k=1}^{n} x_k e_k\right\| \quad \text{where } (x_1, x_2, \ldots, x_n) \text{ are the coordinates of } x \text{ with respect to } \mathscr{B} \\
&\leqslant \sum_{k=1}^{n} |x_k| \cdot \|e_k\| \quad \text{(by the triangle inequality)} \\
&\leqslant M\left(\sum_{k=1}^{n} |x_k|\right) = M\|x\|_1.
\end{aligned}$$

So now it is sufficient to show that there exists a positive real number m such that $m\|x\|_1 \leqslant \|x\|$ for all $x \in V$. This is what we do next. Since

$$\Big|\|x\| - \|y\|\Big| \leqslant \|x - y\| \leqslant M\|x - y\|_1 \quad \text{for all } x, y \in V$$

we see that the mapping $x \mapsto \|x\|$ is continuous on $(V, \|\cdot\|_1)$. Let us now show that $S_1 := \{x \in V : \|x\|_1 = 1\}$ is a compact subset of $(V, \|\cdot\|_1)$. Consider $Y := [-1, 1]^n$ endowed with the product topology and let $D := \{(x_1, x_2, \ldots, x_n) \in Y : \sum_{k=1}^{n} |x_k| = 1\}$. Then D is a closed subset of Y and hence D is compact. Now, $S_1 = \varphi(D)$, where $\varphi : D \to V$ is defined by, $\varphi(x_1, x_2, \ldots, x_n) := \sum_{k=1}^{n} x_k e_k$. However, since $\varphi : D \to (V, \|\cdot\|_1)$ is continuous, S_1 is compact. Hence there exists a point $x_0 \in S_1$ such that $0 < m := \|x_0\| \leqslant \|x\|$ for all $x \in S_1$. Therefore, $m \leqslant \|(x/\|x\|_1)\|$ for any $x \in V \setminus \{0\}$ and so $m\|x\|_1 \leqslant \|x\| \leqslant M\|x\|_1$ for all $x \in V$. $\square$

Corollary 2.9. *Every finite dimensional normed linear space is a Banach space.*

Proof. Let $(X, \|\cdot\|)$ be a finite dimensional normed linear space with basis $\mathscr{B} := \{e_1, e_2, \ldots, e_n\}$. Define the $\|\cdot\|_\infty$ norm on X by, $\|x\|_\infty := \max\{|x_k| : 1 \leqslant k \leqslant n\}$ where $(x_k : 1 \leqslant k \leqslant n)$ are the coordinates of x with respect to $\mathscr{B}$. Then it is easy to check that $(X, \|\cdot\|_\infty)$ is a Banach space. Since the norms $\|\cdot\|$ and $\|\cdot\|_\infty$ are equivalent $(X, \|\cdot\|)$ is also a Banach space. $\square$

Corollary 2.10. *Let $(Y, \|\cdot\|)$ be a finite dimensional subspace of $(X, \|\cdot\|)$. Then Y is a closed subspace of $(X, \|\cdot\|)$.*

Proof. By the previous corollary, $(Y, \|\cdot\|)$ is a Banach space. Hence, if we define a metric $\rho : X^2 \to \mathbb{R}$ by, $\rho(x, y) := \|x - y\|$ for all $x, y \in X$, then $(Y, \rho|_Y)$ is a complete metric space. Therefore, from metric space theory, Y is a closed subset of (X, ρ). This proves the result. $\square$

Theorem 2.11. *Let $(X, \|\cdot\|)$ be a finite dimensional normed linear space. Then B_X is compact in $(X, \|\cdot\|)$.*

Proof. Let $(X, \|\cdot\|)$ be a finite dimensional normed linear space with basis $\mathscr{B} := \{e_1, e_2, \ldots, e_n\}$. Define the $\|\cdot\|_\infty$ norm on X by, $\|x\|_\infty := \max\{|x_k| : 1 \leqslant k \leqslant n\}$, where $(x_k : 1 \leqslant k \leqslant n)$

are the coordinates of x with respect to $\mathscr{B}$. Consider $Y := [-1,1]^n$ endowed with the product topology and let $\varphi : Y \to X$ be defined by, $\varphi(x_1, x_2, \ldots, x_n) := \sum_{k=1}^{n} x_k e_k$. Then $B_1 := \varphi(Y)$ is compact in $(X, \|\cdot\|_\infty)$ since Y is compact and $\varphi : Y \to (X, \|\cdot\|_\infty)$ is continuous. Now B_1 is the closed unit ball in $(X, \|\cdot\|_\infty)$. Hence, there exists a $0 < m < \infty$ such that $mB_X \subseteq B_1$ since $\|\cdot\|_\infty$ and $\|\cdot\|$ are equivalent norms. Moreover, since the norms $\|\cdot\|_\infty$ and $\|\cdot\|$ are equivalent, B_X is closed in $(X, \|\cdot\|_\infty)$. Therefore, mB_X is closed in B_1 and thus compact. It now follows that B_X is compact in $(X, \|\cdot\|_\infty)$ since the mapping $x \mapsto (1/m)x$ is continuous on $(X, \|\cdot\|_\infty)$. Finally, since $\|\cdot\|_\infty$ and $\|\cdot\|$ are equivalent norms, B_X is compact in $(X, \|\cdot\|)$. $\square$

Exercise 2.12. *Let C be a nonempty closed subset of a normed linear space $(X, \|\cdot\|)$ and let $0 < r < 1$. Show that $C = \bigcap_{n\in\mathbb{N}}(C + r^n B_X)$.*

Let T be a subset of a metric space (X, ρ). Then we say that T is **totally bounded** if for every $0 < \varepsilon$ there exists a finite set $F_\varepsilon \subseteq X$ such that $T \subseteq \bigcup\{B[x;\varepsilon] : x \in F_\varepsilon\}$.

Theorem 2.13. *Let $(X, \|\cdot\|)$ be a normed linear space. Then $(X, \|\cdot\|)$ is finite dimensional if, and only if, B_X is totally bounded.*

Proof. If $(X, \|\cdot\|)$ is finite dimensional, then B_X is compact and hence totally bounded. Conversely, suppose that B_X is totally bounded. Fix $0 < r < 1$. Since B_X is totally bounded there exists a finite subset F of X such that $B_X \subseteq \bigcup_{x\in F} B[x;r]$. Let $Y := \mathrm{sp}(F)$. Then Y is finite dimensional and hence a closed subspace of $(X, \|\cdot\|)$ and $B_X \subseteq Y + rB_X$.

We claim that $X = Y$. To see this consider the following argument. Let $n \in \mathbb{N}$, then

$$r^n B_X \subseteq r^n(Y + rB_X) = r^n Y + r^{n+1} B_X = Y + r^{n+1} B_X.$$

Therefore,

$$Y + r^n B_X \subseteq Y + (Y + r^{n+1} B_X) = (Y + Y) + r^{n+1} B_X = Y + r^{n+1} B_X.$$

Thus, by induction, it follows that $B_X \subseteq Y + r^n B_X$ for all $n \in \mathbb{N}$. Hence, by the Exercise 2.12, $B_X \subseteq Y$. This shows that $X = Y$, which in turn means $(X, \|\cdot\|)$ is finite dimensional. This completes the proof. $\square$

Corollary 2.14. *Let $(X, \|\cdot\|)$ be a normed linear space. Then $(X, \|\cdot\|)$ is finite dimensional if, and only if, B_X is compact.*

Next, we consider one of the fundamental building blocks of Banach space theory.

If C is a nonempty subset of a metric space (M, d), then for each $x \in M$,

$$\mathrm{dist}(x, C) := \inf\{d(x, c) : c \in C\}.$$

Exercise 2.15. *Let Y be a proper closed subspace of a normed linear space $(X, \|\cdot\|)$. Show that (i) $\mathrm{dist}(x, Y) = 0$ if, and only if, $x \in Y$; (ii) $\mathrm{dist}(\lambda x, Y) = |\lambda| \mathrm{dist}(x, Y)$ for all $\lambda \in \mathbb{K}$ and $x \in X$; (iii) $\mathrm{dist}(x + y, Y) = \mathrm{dist}(x, Y)$ for all $x \in X$ and $y \in Y$.*

Lemma 2.16 (Riesz's Lemma)**.** *Let Y be a proper closed subspace of a normed linear space $(X, \|\cdot\|)$. Then for every $0 < \varepsilon$ there exists a $z \in S_X$ such that $1-\varepsilon \leqslant \operatorname{dist}(z, Y)$.*

Proof : Choose $x' \notin Y$. Then $\operatorname{dist}(x', Y) > 0$. Next, let us choose $0 < t$ so that $1 - \varepsilon < t \operatorname{dist}(x', Y) < 1$. Set $x := tx'$, then $1 - \varepsilon < \operatorname{dist}(x, Y) < 1$, since

$$t \operatorname{dist}(x', Y) = \operatorname{dist}(tx', Y) = \operatorname{dist}(x, Y).$$

Pick any $y \in Y$ such that $\|x - y\| \leqslant 1$ and set $z := (x - y)/\|x - y\|$. Then $z \in S_X$ and

$$\begin{aligned} 1 - \varepsilon < \operatorname{dist}(x, Y) &\leqslant (1/\|x - y\|) \operatorname{dist}(x, Y) \\ &= (1/\|x - y\|) \operatorname{dist}(x - y, Y) \\ &= \operatorname{dist}((x - y)/\|x - y\|, Y). \end{aligned}$$

This completes the proof. $\square$.

Exercise 2.17. *Let T be a subset of a metric space (X, ρ). Show that T is not totally bounded if, and only if, there exists an $0 < \varepsilon$ and an infinite subset C of T such that $\varepsilon < \rho(x, y)$ for all $(x, y) \in C^2 \setminus \Delta_C$.*

We now give a second proof of the following fact.

Theorem 2.18. *Let $(X, \|\cdot\|)$ be a normed linear space. If $(X, \|\cdot\|)$ is infinite dimensional, then B_X is not totally bounded.*

Proof. If $(X, \|\cdot\|)$ is infinite dimensional, then by Riesz's Lemma we can inductively construct a sequence $(x_n : n \in \mathbb{N})$ in S_X such that $1/2 < \operatorname{dist}(x_{n+1}, \operatorname{span}\{x_1, x_2, \ldots, x_n\})$. Thus, $1/2 < \|x_m - x_n\|$ whenever $m \neq n$. Therefore, B_X is not totally bounded. $\square$

Linear Operators

We call a subset A of a normed linear space $(X, \|\cdot\|)$ **bounded** if there exists an $r \in [0, \infty)$ such that $A \subseteq rB_X$. Let $(X, \|\cdot\|)$ and $(Y, |||\cdot|||)$ be normed linear spaces and let $T : X \to Y$ be a linear mapping. Then we say that T is a **bounded** linear mapping if $T(B_X)$ is a bounded subset of Y. For a bounded linear mapping T acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, |||\cdot|||)$ we define the **operator norm** of T to be,

$$\|T\| := \sup\{|||T(x)||| : x \in B_X\}.$$

Exercise 2.19. *Let T be a bounded linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, |||\cdot|||)$. Show that $\|T\| = \sup_{x \in B_X \setminus \{0\}} \dfrac{|||T(x)|||}{\|x\|} = \sup_{x \in S_X} |||T(x)|||$.*

Note: $|||T(x)||| \leqslant \|T\|\|x\|$ for all $x \in X$. In fact, $\|T\|$ is the smallest real number M such that $|||T(x)||| \leqslant M\|x\|$ for all $x \in X$.

Theorem 2.20. *Let $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$ be normed linear spaces and let $T : X \to Y$ be a linear mapping. Then the following properties are equivalent:*

(i) *T is a bounded operator;*

(ii) *T is continuous at 0;*

(iii) *T is continuous on X.*

Proof. *(i)* $\Rightarrow$ *(ii)*: Suppose that T is a bounded operator. Then there exists a $K > 0$ such that $|\!|\!|T(x)|\!|\!| \leqslant K\|x\|$ for all $x \in X$. (Note: we could take $K = \|T\|$). Suppose $\varepsilon > 0$ is given. Let $\delta := \varepsilon/K > 0$. Then $|\!|\!|T(x) - T(0)|\!|\!| = |\!|\!|T(x)|\!|\!| \leqslant K\|x\| = K\|x - 0\| < \varepsilon$ for all $\|x - 0\| < \delta$. This shows that T is continuous at $x = 0$.

(ii) $\Rightarrow$ *(i)*: Suppose that T is continuous at 0. Let $\varepsilon := 1$. Then there exists a $\delta > 0$ such that

$$\delta T(B_X) = T(\delta B_X) = T(B[0;\delta]) \subseteq B[T(0);\varepsilon] = B[0;\varepsilon] = \varepsilon B_Y = B_Y.$$

Therefore, $T(B_X) \subseteq (1/\delta)B_Y$ and so T is bounded.

(i) $\Rightarrow$ *(iii)*: Suppose that T is bounded. Then there exists a $K > 0$ such that $|\!|\!|T(x)|\!|\!| \leqslant K\|x\|$ for all $x \in X$. Now suppose that $x_0 \in X$ and $\varepsilon > 0$ are given. Let $\delta := \varepsilon/K$. Then,

$$|\!|\!|T(x) - T(x_0)|\!|\!| = |\!|\!|T(x - x_0)|\!|\!| \leqslant K\|x - x_0\| < \varepsilon.$$

for all $\|x - x_0\| < \delta$.

(iii) $\Rightarrow$ *(ii)*: This is obvious. $\square$

Let $(X, \|\cdot\|)$ and $(Y, |\!|\!|\cdot|\!|\!|)$ be normed linear spaces (over a field $\mathbb{K}$). Then by $B(X,Y)$ we denote the space of all bounded linear operators from X into Y. It is easy to show that $B(X,Y)$ is a vector space (over $\mathbb{K}$).

Theorem 2.21. *Let $(X, \|\cdot\|)$ and $(Y, |\!|\!|\cdot|\!|\!|)$ be normed linear spaces. Then $B(X,Y)$, equipped with the operator norm, is a normed linear space.*

Proof. We need only show that the "operator norm" is indeed a norm. Let $T \in B(X,Y)$, then $\|T\| = \sup_{x \in S_X} |\!|\!|T(x)|\!|\!|$. Hence, $\|T\| \geqslant 0$ and $\|T\| = 0$ if, and only if, $T = 0$. Now, let $\lambda \in \mathbb{K}$ and $T \in B(X,Y)$, then

$$\|\lambda T\| = \sup_{x \in S_X} |\!|\!|(\lambda T)(x)|\!|\!| = \sup_{x \in S_X} |\lambda| \cdot |\!|\!|T(x)|\!|\!| = |\lambda| \sup_{x \in S_X} |\!|\!|T(x)|\!|\!| = |\lambda| \cdot \|T\|.$$

Finally, if $S, T \in B(X,Y)$, then for any $x \in S_X$,

$$|\!|\!|(S + T)(x)|\!|\!| \leqslant |\!|\!|S(x)|\!|\!| + |\!|\!|T(x)|\!|\!| \leqslant \|S\| + \|T\|.$$

Therefore, $\|S + T\| = \sup_{x \in S_X} |\!|\!|(S + T)(x)|\!|\!| \leqslant \|S\| + \|T\|$. $\square$

Let $(X, \|\cdot\|)$ be a normed linear space. Then we shall denote by X^* the vector space of all bounded linear functionals on X. The space X^* equipped with the operator norm is called the **dual space** of X and is a normed linear space since $X^* = B(X, \mathbb{K})$. The norm on X^* is usually called the **dual norm** (on X^*) instead of the "operator norm".

Theorem 2.22. *Let $(X, \|\cdot\|)$ be a normed linear space and let $(Y, \|\cdot\|)$ be a Banach space. Then $B(X, Y)$ is a Banach space.*

Proof. Let $(T_n : n \in \mathbb{N})$ be a Cauchy sequence in $B(X, Y)$. Then for each $x \in X$, $(T_n(x) : n \in \mathbb{N})$ is a Cauchy sequence in $(Y, \|\cdot\|)$ since,

$$\|T_n(x) - T_m(x)\| = \|(T_n - T_m)(x)\| \leqslant \|T_n - T_m\| \cdot \|x\|.$$

Since $(Y, \|\cdot\|)$ is complete the sequence $(T_n(x) : n \in \mathbb{N})$ is convergent in $(Y, \|\cdot\|)$. For each $x \in X$, let $T(x) := \lim_{n\to\infty} T_n(x)$. Then $T : X \to Y$ is well-defined and linear. Since $(T_n : n \in \mathbb{N})$ is a Cauchy sequence in $B(X, Y)$, it is bounded in $B(X, Y)$, i.e., there exists a constant $M > 0$ such that $\|T_n\| \leqslant M$ for all $n \in \mathbb{N}$. We claim that $\|T\| \leqslant M$. Let $x \in S_X$, then

$$\|T(x)\| = \left\|\lim_{n\to\infty} T_n(x)\right\| = \lim_{n\to\infty} \|T_n(x)\| \leqslant \sup_{n\in\mathbb{N}} \|T_n(x)\| \leqslant \sup_{n\in\mathbb{N}} \|T_n\| \leqslant M.$$

Therefore, $\|T\| \leqslant M$. We now claim that $(T_n : n \in \mathbb{N})$ converges to T with respect to the operator norm on $B(X, Y)$. To justify this claim let us consider an arbitrary $\varepsilon > 0$. Then there exists a $N \in \mathbb{N}$ such that

$$\|T_m(x) - T_n(x)\| \leqslant \|T_m - T_n\| < \varepsilon \quad \text{for all } x \in B_X \text{ and all } m, n > N.$$

Thus, if we take the limit over $m \in \mathbb{N}$ we get that

$$\|(T - T_n)(x)\| = \|T(x) - T_n(x)\| \leqslant \varepsilon \quad \text{for all } x \in B_X \text{ and all } n > N.$$

Hence, we have that $\|T - T_n\| = \sup\{\|(T - T_n)(x)\| : x \in B_X\} \leqslant \varepsilon$ for all $n > N$. $\square$

Theorem 2.23. *All linear operators defined on finite dimensional normed linear spaces are continuous.*

Proof. Let $(X, \|\cdot\|)$ be a finite dimensional normed linear space, $(Y, \|\cdot\|)$ be a normed linear space and $T : X \to Y$ be a linear operator. Let us define a norm $|\!|\!|\cdot|\!|\!|$ by, $|\!|\!|x|\!|\!| := \|x\| + \|T(x)\|$ for all $x \in X$. By the Fundamental Theorem of Finite Dimensional Normed Linear Spaces, there exists a constant $M > 0$ such that $|\!|\!|x|\!|\!| \leqslant M\|x\|$ for all $x \in X$. This implies that $\|T(x)\| \leqslant M\|x\|$ for all $x \in X$, i.e., $T \in B(X, Y)$. $\square$

A linear transformation $T : (X, \|\cdot\|) \to (Y, |\!|\!|\cdot|\!|\!|)$ acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, |\!|\!|\cdot|\!|\!|)$ is called a **normed linear space isomorphism** if:

(i) T is one-to-one and onto;

(ii) $T \in B(X, Y)$;

(iii) $T^{-1} \in B(Y, X)$.

If there exists an isomorphism T acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, |||\cdot|||)$, then we say that $(X, \|\cdot\|)$ is **isomorphic to** $(Y, |||\cdot|||)$

Corollary 2.24. *Any two n-dimensional normed linear spaces (over the same field $\mathbb{K}$) are isomorphic.*

Proof. Suppose that $(X, \|\cdot\|$ and $(Y, |||\cdot|||)$ are n-dimensional normed linear spaces. Let $\mathscr{I} : X \to Y$ be any vector space isomorphism from X into Y. Note that such an isomorphism exists since X and Y have the same dimension. Since $\mathscr{I}$ is one-to-one and onto, $\mathscr{I}^{-1} : Y \to X$ exists. Moreover, $\mathscr{I}^{-1}$ will also be linear. The result now follows from Theorem 2.23. □

Exercise 2.25. *Show that a normed linear space $(X, \|\cdot\|)$ is finite dimensional if, and only if, every linear functional on $(X, \|\cdot\|)$ is continuous.*

These last two results indicate that the isomorphic theory of finite dimensional normed linear spaces largely reduces to linear algebra.

Chapter 3

Hilbert Spaces

Recall that an **inner product** (or a scalar product or a dot product) on a vector space X is a scalar-valued function $\langle \cdot, \cdot \rangle$ on $X \times X$ such that:

(i) for every $y \in X$, the function $x \mapsto \langle x, y \rangle$ is linear;

(ii) $\overline{\langle x, y \rangle} = \langle y, x \rangle$ for every $x, y \in X$;

(iii) $\langle x, x \rangle \geqslant 0$ for every $x \in X$;

(iv) $\langle x, x \rangle = 0$ if, and only if, $x = 0$.

Note that by (i), $\langle 0, y \rangle = 0$ for any $y \in X$, and so by (ii), $\langle y, 0 \rangle = \overline{0} = 0$.

Theorem 3.1 (Cauchy-Schwarz inequality)**.** *Let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space X.*

(i) *For any $x, y \in X$, we have $|\langle x, y \rangle| \leqslant \sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle}$;*

(ii) *the function $\|x\| := \sqrt{\langle x, x \rangle}$ is a norm on X.*

Proof. (i): If $\langle y, y \rangle = 0$, then we have that $y = 0$ and the inequality is satisfied. So we may suppose that $\langle y, y \rangle > 0$. Then for any $\lambda \in \mathbb{K}$,

$$\begin{aligned} 0 \leqslant \|x - \lambda y\|^2 &= \langle x - \lambda y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \lambda \langle y, x \rangle - \overline{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle \\ &= \langle y, y \rangle \left[\left| \lambda - \frac{\langle x, y \rangle}{\langle y, y \rangle} \right|^2 + \left[\frac{\langle x, x \rangle}{\langle y, y \rangle} - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle^2} \right] \right]. \end{aligned}$$

Set $\lambda := \langle x, y \rangle / \langle y, y \rangle$ and multiply both sides by $\langle y, y \rangle$. Then,

$$|\langle x, y \rangle|^2 \leqslant \langle x, x \rangle \langle y, y \rangle.$$

(ii): We will check the triangle inequality. For any $x, y \in X$, we have

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle y, y \rangle + \langle x, y \rangle + \langle y, x \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + 2\mathrm{Real}\langle x, y \rangle \leqslant \langle x, x \rangle + \langle y, y \rangle + 2|\langle x, y \rangle| \\ &\leqslant \langle x, x \rangle + \langle y, y \rangle + 2\sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle} \\ &= (\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle})^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

This concludes the proof $\square$

Exercise 3.2. *Show that* $|\langle x,y\rangle| = \sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}$ *if, and only if,* x *and* y *are linearly dependent.*

One immediate consequence of Theorem 3.1 is that $\langle\cdot,\cdot\rangle$ is a continuous function on $(X\|\cdot\|)\times(X,\|\cdot\|)$ into the scalar field. In particular, it implies that for a fixed vector $y \in X$, $x \mapsto \langle x,y\rangle$ is a continuous linear functional on X.

An ordered pair $(H,\langle\cdot,\cdot\rangle)$ is called a **Hilbert space** if:

(i) H is a vector space;

(ii) $\langle\cdot,\cdot\rangle$ is an inner product on H and

(iii) $(H,\|\cdot\|)$ is a Banach space, where $\|x\|^2 = \langle x,x\rangle$ for all $x \in H$.

Theorem 3.3. *Let* $(V,\|\cdot\|)$ *be a normed linear space. Then there exists an inner product* $\langle\cdot,\cdot\rangle\colon V\times V \to \mathbb{K}$ *such that* $\|x\|^2 = \langle x,x\rangle$ *for all* $x \in V$ *if, and only if, the norm* $\|\cdot\|$ *satisfies the parallelogram law, i.e.,*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \textit{for all } x,y \in V.$$

Moreover, the inner product $\langle\cdot,\cdot\rangle$ *is generated by the polarisation identity*

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \quad \textit{for all } x,y\in V,\ \textit{if } V \textit{ is a vector space over } \mathbb{R}$$

and by

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) \quad \textit{for all } x,y\in V$$

if V *is a vector space over* $\mathbb{C}$*. Alternatively, we can write* $\langle x,y\rangle = \frac{1}{4}\sum_{k=1}^{4} i^k\|x+i^k y\|^2$.

Proof. $(\Rightarrow)$ Suppose that the norm $\|\cdot\|$ is induced by the inner product $\langle\cdot,\cdot\rangle$. Then

$$\begin{aligned}\|x+y\|^2 &= \langle x+y,x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \quad \text{and}\\ \|x-y\|^2 &= \langle x-y,x-y\rangle = \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle.\end{aligned}$$

Taking the sum gives the parallelogram law:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\langle x,x\rangle + \langle y,y\rangle) = 2(\|x\|^2 + \|y\|^2).$$

Taking the difference gives:

$$\|x+y\|^2 - \|x-y\|^2 = 2(\langle x,y\rangle + \langle y,x\rangle) = 2(\langle x,y\rangle + \overline{\langle x,y\rangle}) = 4\mathrm{Real}\langle x,y\rangle$$

which is the real part of the polarisation identity. Now,

$$\mathrm{Im}\langle x,y\rangle = \mathrm{Real}(-i\langle x,y\rangle) = \mathrm{Real}\langle x,iy\rangle = \frac{1}{4}\left(\|x+iy\|^2 - \|x-iy\|^2\right).$$

($\Leftarrow$) Suppose the norm satisfies the parallelogram law. It suffices to show that $\langle \cdot, \cdot \rangle$ is a complex inner product if it is defined by the polarisation identity. The proof for real inner product is similar by removing all the imaginary terms.

First we check that $\langle x, y \rangle = \overline{\langle y, x \rangle}$:

$$\begin{aligned}
\langle x, y \rangle &= \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) + \frac{i}{4}(\|x+iy\|^2 - \|x-iy\|^2) \\
&= \frac{1}{4}(\|y+x\|\|^2 - \|y-x\|^2) + \frac{i}{4}(\|(-i)(x+iy)\|^2 - \|i(x-iy)\|^2) \\
&= \frac{1}{4}(\|y+x\|^2 - \|y-x\|^2) - \frac{i}{4}(\|y+ix\|^2 - \|y-ix\|^2) = \overline{\langle y, x \rangle}.
\end{aligned}$$

It follows that $\langle x, x \rangle$ is real, so that we may check $0 \leqslant \langle x, x \rangle$:

$$\langle x, x \rangle = \mathrm{Real}\langle x, x \rangle = \frac{1}{4}(\|x+x\|^2 + \|x-x\|^2) = \|x\|^2 \geqslant 0,$$

and $\langle x, x \rangle = 0$ if, and only if, $\|x\|^2 = 0$, or $x = 0$.

We now show additive distributivity. For $x, y, z \in V$: We will use the identity that $x + i^k y + i^k z = [(1/2)x + i^k y] + [(1/2)x + i^k z]$ and the parallelogram identity.

$$\begin{aligned}
\langle x, y+z \rangle &= \sum_{k=1}^{4} i^k \|x + i^k y + i^k z\|^2 \\
&= \sum_{k=1}^{4} i^k \left(2\left\| \frac{x}{2} + i^k y \right\|^2 + 2\left\| \frac{x}{2} + i^k z \right\|^2 - \|i^k(y-z)\|^2 \right) \\
&= \sum_{k=1}^{4} i^k \left(2\left\| \frac{x}{2} + i^k y \right\|^2 + 2\left\| \frac{x}{2} + i^k z \right\|^2 - \|(y-z)\|^2 \right) \\
&= 2\sum_{k=1}^{4} i^k \left\| \frac{x}{2} + i^k y \right\|^2 + 2\sum_{k=1}^{4} i^k \left\| \frac{x}{2} + i^k z \right\|^2 = 2\left(\left\langle \frac{x}{2}, y \right\rangle + \left\langle \frac{x}{2}, z \right\rangle \right).
\end{aligned}$$

We used the fact that $\sum_{k=1}^{4} i^k c = 0$ for all $c \in \mathbb{C}$. Putting $z = 0$ gives $\langle x, y \rangle = 2\left\langle \frac{x}{2}, y \right\rangle$ so that $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$. We now show scalar multiplication distribution. Using $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$ we can show, by induction, that $\langle ax, y \rangle = a\langle x, y \rangle$ for all $a \in \mathbb{Q}$. The equation then holds for all $a \in \mathbb{R}$ by the continuity of $\|\cdot\|$ and the density of $\mathbb{Q}$ in $(\mathbb{R}, |\cdot|)$. Finally, for complex multiples we have

$$\langle ix, y \rangle = \sum_{k=1}^{4} i^k \|ix + i^k y\|^2 = i \sum_{k=1}^{4} i^{k-1} \|x + i^{k-1} y\|^2 = i\langle x, y \rangle$$

so that $\langle ax, y \rangle = a\langle x, y \rangle$ for all $a \in \mathbb{C}$ by real linearity. This allows us to conclude that $\langle \cdot, \cdot \rangle$ is an inner product on V that induces the norm $\|\cdot\|$. $\square$

Therefore, a Banach space $(X, \|\cdot\|)$ is a Hilbert space if, and only if, every two dimensional subspace of $(X, \|\cdot\|)$ is a Hilbert space.

Let $(H, \langle \cdot, \cdot \rangle)$ be a Hilbert space and let $x, y \in H$. We say that x is **orthogonal** to y, denoted $x \perp y$, if $\langle x, y \rangle = 0$. Let M be a subset of H. We say that $x \in H$ is **orthogonal** to M, denoted $x \perp M$, if x is orthogonal to every vector $y \in M$.

Let M be a subset of a Hilbert space $(H, \langle \cdot, \cdot \rangle)$. Then the set

$$M^{\perp} := \{h \in H : h \perp M\}$$

is called the **orthogonal complement of** M in H.

Exercise 3.4. *Let M be a subspace of a Hilbert space H. Show that (i) $M^{\perp}$ is a closed subspace of H, (ii) $M \cap M^{\perp} = \{0\}$ and (iii) $M \subseteq (M^{\perp})^{\perp}$.*

Lemma 3.5. *Let M be a closed subspace of a Hilbert space $(H, \langle \cdot, \cdot \rangle)$. If $x_0 \in H$, then there exists an $m_0 \in M$ such that $\|x_0 - m_0\| = \inf\{\|x_0 - m\| : m \in M\}$.*

Proof. Choose a sequence $(m_n : n \in \mathbb{N})$ in M such that

$$d := \lim_{n \to \infty} \|x_0 - m_n\| = \inf\{\|x_0 - m\| : m \in M\}.$$

Recall the parallelogram law; namely, $\|x - y\|^2 = 2[\|x\|^2 + \|y\|^2] - \|x + y\|^2$. Let us apply this with $x := (x_0 - m_n)$ and $y := (x_0 - m_m)$, then

$$\begin{aligned}
\|m_m - m_n\|^2 &= 2[\|x_0 - m_n\|^2 + \|x_0 - m_m\|^2] - \|2x_0 - (m_n + m_m)\|^2 \\
&= 2[\|x_0 - m_n\|^2 + \|x_0 - m_m\|^2 - 2\|x_0 - (m_n + m_m)/2\|^2] \\
&\leqslant 2[\|x_0 - m_n\|^2 + \|x_0 - m_m\|^2 - 2d^2], \quad \text{since } (m_n + m_m)/2 \in M.
\end{aligned}$$

It now follows that $(m_n : n \in \mathbb{N})$ is a Cauchy sequence in M. Let $m_0 := \lim_{n \to \infty} m_n$. Then $m_0 \in M$, since M is closed and $\|x_0 - m_0\| = \lim_{n \to \infty} \|x_0 - m_n\| = d$. $\square$

Lemma 3.6. *Let M be a closed subspace of a Hilbert space X. If $x_0 \notin M$ and there exists an $m_0 \in M$ such that $\|x_0 - m_0\| = \inf\{\|x_0 - m\| : m \in M\}$, then $(x_0 - m_0) \in M^{\perp}$.*

Proof. Fix $m \in M$ and define $D : \mathbb{R} \to \mathbb{R}$ by,

$$D(\lambda) := \|x_0 - (m_0 + \lambda m)\|^2 = \|(x_0 - m_0) - \lambda m\|^2.$$

Therefore,

$$D(\lambda) = \lambda^2 \|m\|^2 - 2\lambda \mathrm{Real}\langle m, x_0 - m_0 \rangle + \|x_0 - m_0\|^2; \text{ which is a quadratic in } \lambda.$$

Now, by assumption, D attains its minimum value at $\lambda = 0$ and so by elementary calculus, $0 = D'(0) = 2\mathrm{Real}\langle m, x_0 - m_0 \rangle$, since $D'(\lambda) = 2\lambda \|m\|^2 - 2\mathrm{Real}\langle m, x_0 - m_0 \rangle$.

Thus, for any $m \in M$, $\mathrm{Real}\langle m, x_0 - m_0 \rangle = 0$. Now, if $m \in M$, then im is also in M and so $0 = \mathrm{Real}\langle im, x_0 - m_0 \rangle = -\mathrm{Im}\langle m, x_0 - m_0 \rangle$, i.e., $\mathrm{Im}\langle m, x_0 - m_0 \rangle = 0$ and so $\langle m, x_0 - m_0 \rangle = 0$. Since $m \in M$ was arbitrary it follows that $(x_0 - m_0) \in M^{\perp}$. $\square$

Theorem 3.7. *If M is a closed subspace of a Hilbert space $(H, \langle\cdot,\cdot\rangle)$, then $M + M^{\perp} = H$. In fact, $M \oplus M^{\perp} = H$.*

Proof. Clearly, $M + M^{\perp} \subseteq H$. So it is sufficient to show that $H \subseteq M + M^{\perp}$. Let $x_0 \in H$, then by the earlier two lemmas there exists a $m_0 \in M$ such that $(x_0 - m_0) \in M^{\perp}$. Thus, $x_0 = m_0 + (x_0 - m_0) \in M + M^{\perp}$. $\square$

Corollary 3.8. *If M is a closed subspace of a Hilbert space $(H, \langle\cdot,\cdot\rangle)$, then $(M^{\perp})^{\perp} = M$.*

Proof. From before we know that $M \subseteq (M^{\perp})^{\perp}$ so it is sufficient to show that $(M^{\perp})^{\perp} \subseteq M$. To this end, choose $x \in (M^{\perp})^{\perp}$. Then $x = m + m^{\perp}$ for some $m \in M$ and $m^{\perp} \in M^{\perp}$ (as $H = M \oplus M^{\perp}$). Now, since $x \in (M^{\perp})^{\perp}$,

$$0 = \langle x, m^{\perp}\rangle = \langle m + m^{\perp}, m^{\perp}\rangle = \langle m, m^{\perp}\rangle + \langle m^{\perp}, m^{\perp}\rangle = 0 + \|m^{\perp}\|^2.$$

Hence $m^{\perp} = 0$ and so $x = m \in M$. $\square$

Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $S \subseteq H$. Then S is called an **orthonormal set** if $\langle s, s'\rangle = 0$ whenever $s \neq s'$ and $\langle s, s\rangle = 1$ for every $s \in S$. A subset S of a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is called an **orthonormal basis** for H if S is an orthonormal set and $H = \overline{\mathrm{span}}(S)$.

Theorem 3.9. *Every nonzero Hilbert space admits an orthonormal basis.*

Proof. Let $(H, \langle\cdot,\cdot\rangle)$ be a nonzero Hilbert space and let X be the family of all orthonormal subsets of H. Then $X \neq \varnothing$ and $(X, \subseteq)$ is a partially ordered set. (Note: if $x \in S_H$, then $\{x\} \in X$). We claim that X contains a maximal element. By Zorn's Lemma to show this we need only show that every totally ordered subset of X has an upper bound. Let $\varnothing \neq T \subseteq X$ be a totally ordered and let $B := \bigcup\{S : S \in T\}$. Clearly, $S \subseteq B$ for each $S \in T$ and so B is an upper bound for T provided we have $B \in X$. So suppose that $x, y \in B$ and $x \neq y$. Then there exists $S_x \in T$ and $S_y \in T$ such that $x \in S_x$ and $y \in S_y$. Now since T is totally ordered either $S_x \subseteq S_y$ or $S_y \subseteq S_x$. Therefore, either $\{x, y\} \subseteq S_x$ or $\{x, y\} \subseteq S_y$. Hence, in either case, $\langle x, y\rangle = 0$. Furthermore, it is easy to see that if $x \in B$, then $\|x\| = 1$. This shows that $B \in X$. Let $B_{\max}$ be a maximal element of $(X, \subseteq)$. We claim that $\overline{\mathrm{span}}(B_{\max}) = H$; for if this is not the case then we may choose $x \in S_H \cap \overline{\mathrm{span}}(B_{\max})^{\perp}$ and set $B^* := B_{\max} \cup \{x\}$. Then $B^* \in X$, $B_{\max} \subseteq B^*$, but $B^* \neq B_{\max}$; which contradicts that maximality of $B_{\max}$. Hence $B_{\max}$ is an orthonormal basis for H. $\square$

Exercise 3.10 (Pythagoras' Theorem)**.** *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. Show that if $x \perp y$, then*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Theorem 3.11. *Let $\{e_k : 1 \leqslant k \leqslant n\}$ be an orthonormal set in an inner product space $(X, \langle\cdot,\cdot\rangle)$. Let $x \in X$ and $M := \mathrm{span}\{e_1, e_2, \ldots e_n\}$. Then:*

(i) $(x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k) \perp M$;

(ii) $\sum_{k=1}^{n}\langle x, e_k\rangle e_k$ *is the closest point in* M *to* x*;*

(iii) $\|x\|^2 = \|\sum_{k=1}^{n}\langle x, e_k\rangle e_k\|^2 + \|x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k\|^2$.

Proof. (i): To show this it is sufficient to check that $\langle x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k, e_j\rangle = 0$ for each $1 \leqslant j \leqslant n$. But this follows from the following simple calculation.

$$\begin{aligned}
\langle x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k, e_j\rangle &= \langle x, e_j\rangle - \langle \sum_{k=1}^{n}\langle x, e_k\rangle e_k, e_j\rangle \\
&= \langle x, e_j\rangle - \sum_{k=1}^{n}\langle x, e_k\rangle\langle e_k, e_j\rangle \\
&= \langle x, e_j\rangle - \sum_{k=1}^{n}\langle x, e_k\rangle\delta_{k,j} = \langle x, e_j\rangle - \langle x, e_j\rangle = 0.
\end{aligned}$$

(ii): Let $m \in M$. Then $m = \sum_{k=1}^{n} m_k e_k$ for some $(m_1, m_2, \dots, m_n) \in \mathbb{K}^n$. Now,

$$\|x - m\|^2 = \left\|(x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k) + \sum_{k=1}^{n}(\langle x, e_k\rangle - m_k)e_k\right\|^2.$$

Therefore, since $\sum_{k=1}^{n}(\langle x, e_k\rangle - m_k)e_k \in M$ and $x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k \in M^{\perp}$ we have that

$$\|x - m\|^2 = \left\|\sum_{k=1}^{n}(\langle x, e_k\rangle - m_k)e_k\right\|^2 + \left\|x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k\right\|^2 \geqslant \left\|x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k\right\|^2$$

i.e., $\|x - m\| \geqslant \|x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k\|$. (iii): The proof of this follows from part (i) and Exercise 3.10. $\square$

Exercise 3.12. *Let* (M, d) *be a metric space. Show that* (M, d) *is not separable if, and only if, there exists an* $\varepsilon > 0$ *and an uncountable set* $C \subseteq M$ *such that* $d(x, y) > \varepsilon$ *for all* $(x, y) \in C^2 \setminus \Delta$*. Here* $\Delta := \{(x, y) \in C^2 : x = y\}$ *- the diagonal of* C^2*.*

Theorem 3.13. *Let* $(H, \langle\cdot,\cdot\rangle)$ *be a separable infinite dimensional Hilbert space. Then* $(H, \langle\cdot,\cdot\rangle)$ *has an orthonormal basis* $\{e_n : n \in \mathbb{N}\}$ *such that* $x = \sum_{k=1}^{\infty}\langle x, e_k\rangle e_k$*, for each* $x \in H$*.*

Proof. We know, from Theorem 3.9 that $(H, \langle\cdot,\cdot\rangle)$ has an orthonormal basis B. Since H is infinite dimensional, B must be infinite. On the other hand, for every $(b, b') \in B^2 \setminus \Delta$, $\|b - b'\|^2 = \|b\|^2 + \|b'\|^2 = 2$ and so by Exercise 3.12, B must be at most countable i.e., B can be expressed as $B = \{e_n : n \in \mathbb{N}\}$. For each $n \in \mathbb{N}$, let $M_n := \operatorname{span}\{e_1, e_2, \dots e_n\}$. Then $\operatorname{span}(B) = \bigcup_{n\in\mathbb{N}} M_n$. Fix $x \in H$. Since $M_n \subseteq M_{n+1}$ for all $n \in \mathbb{N}$, $0 \leqslant \operatorname{dist}(x, M_{n+1}) \leqslant \operatorname{dist}(x, M_n)$ and so $\lim_{n\to\infty}\operatorname{dist}(x, M_n)$ exists, and is greater than, or equal to 0. Further, since $H = \overline{\operatorname{span}}(B) = \overline{\bigcup_{n\in\mathbb{N}} M_n}$, it follows that $\lim\limits_{n\to\infty}\operatorname{dist}(x, M_n) = 0$. However, by Theorem 3.11 part (ii)

$$\operatorname{dist}(x, M_n) = \left\|x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k\right\| \quad \text{for each } n \in \mathbb{N}.$$

Thus, $\lim\limits_{n\to\infty}\|x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k\| = 0$ and we are done. $\square$

Exercise 3.14. *Let $\{e_k : 1 \leqslant k \leqslant n\}$ be an orthonormal set in an inner product space $(X, \langle\cdot,\cdot\rangle)$. Show that for any $(a_1, a_2, \ldots a_n) \in \mathbb{K}^n$, $\|\sum_{k=1}^n a_k e_k\|^2 = \sum_{k=1}^n |a_k|^2$.*

Theorem 3.15. *Let $\{e_n : n \in \mathbb{N}\}$ be an orthonormal set in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ and let $x \in H$. Then: (i) $\sum_{k=1}^\infty |\langle x, e_k\rangle|^2 \leqslant \|x\|^2$ (Bessel's Inequality); (ii) If $\{e_n : n \in \mathbb{N}\}$ is an orthonormal basis for $(H, \langle\cdot,\cdot\rangle)$, then $\sum_{k=1}^\infty |\langle x, e_k\rangle|^2 = \|x\|^2$ (Parseval's Identity).*

Proof. (i): For every $n \in \mathbb{N}$, we have

$$0 \leqslant \left\| x - \sum_{k=1}^n \langle x, e_k\rangle e_k \right\|^2 = \|x\|^2 - \left\| \sum_{k=1}^n \langle x, e_k\rangle e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x, e_k\rangle|^2.$$

From which Bessel's inequality follows.

(ii): If $\{e_n : n \in \mathbb{N}\}$ is an orthonormal basis for $(H, \langle\cdot,\cdot\rangle)$, then $x = \sum_{k=1}^\infty \langle x, e_k\rangle e_k$. The result then follows from the above equation. $\Box$

Example 3.16. *Recall that $\ell^2(\mathbb{N}) := \{(x_n : n \in \mathbb{N}) \in \mathbb{K}^{\mathbb{N}} : \sum_{n\in\mathbb{N}} |x_n|^2 < \infty\}$. On $\ell^2(\mathbb{N})$ one can define the following inner product.*

$$\langle (x_n : n \in \mathbb{N}), (y_n : n \in \mathbb{N})\rangle_2 := \sum_{n\in\mathbb{N}} x_n \overline{y_n}.$$

Then $(\ell^2(\mathbb{N}), \langle\cdot,\cdot\rangle_2)$ is a separable infinite dimensional Hilbert space.

We now present a representation theorem for separable infinite dimensional Hilbert spaces.

Theorem 3.17 (Riesz-Fischer Theorem)**.** *Every separable infinite dimensional Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is isometrically isomorphic to $(\ell^2(\mathbb{N}), \langle\cdot,\cdot\rangle_2)$.*

Proof : Let $\{e_n : n \in \mathbb{N}\}$ be an orthonormal basis for $(H, \langle\cdot,\cdot\rangle)$. Define $T : \ell^2(\mathbb{N}) \to H$ by, $T(a) := \sum_{k=1}^\infty a_k e_k$, where $a := (a_k : k \in \mathbb{N})$. First we must show that T is well-defined, i.e., show that for each $a \in \ell^2(\mathbb{N})$, $T(a)$ really is an element of H. Let $a := (a_k : k \in \mathbb{N}) \in \ell^2(\mathbb{N})$, then for each $(m, n) \in \mathbb{N}^2$ such that $m < n$ we have that $\|\sum_{k=m}^n a_k e_k\|^2 = \sum_{k=m}^n |a_k|^2$. Therefore, the partial sums $(\sum_{k=1}^n a_k e_k : n \in \mathbb{N})$ form a Cauchy sequence in $(H, \langle\cdot,\cdot\rangle)$ and thus are convergent. It is easy to see that T is linear and by Parseval's Identity it follows that T is an isometric embedding. Therefore, it remains to show that T is onto. To this end, consider $x \in H$. Then $x = \sum_{k=1}^\infty \langle x, e_k\rangle e_k$. Define $a := (a_k : k \in \mathbb{N})$ by, $a_k := \langle x, e_k\rangle$. By Bessel's inequality $a \in \ell^2(\mathbb{N})$. The proof is completed with the simple observation that $T(a) = x$. $\Box$

Example 3.18. *Let Γ be a nonempty set and let $p \in [1, \infty)$. We shall denote by, $\ell^p(\Gamma)$ the set of all functions from Γ into $\mathbb{K}$ such that*

$$\sup\left\{ \sum_{\gamma\in F} |f(\gamma)|^p : F \textit{ is a finite subset of } \Gamma \right\} < \infty.$$

Then $(\ell^p(\Gamma), \|\cdot\|_p)$ *is a Banach space, where*

$$\|f\|_p := \left(\sup\left\{\sum_{\gamma\in F}|f(\gamma)|^p : F \textit{ is a finite subset of } \Gamma\right\}\right)^{1/p}.$$

Note that if $f \in \ell^p(\Gamma)$*, then* $\{\gamma \in \Gamma : f(\gamma) \neq 0\}$ *is at most countable and we write* $\|f\|_p = (\sum_{\gamma\in\Gamma}|f(\gamma)|^p)^{1/p}$*. If* $p = 2$*, then* $\ell^p(\Gamma)$ *is a Hilbert space with inner product defined by*

$$\langle f, g\rangle_2 := \sum_{\gamma\in\Gamma} f(\gamma)\overline{g(\gamma)}.$$

If $\Gamma = \{1, 2, \ldots, n\}$*, then we write* ℓ^2_n *instead of* $\ell^2(\{1, 2, \ldots, n\})$*.*

Exercise 3.19. *Let* $(H, \langle\cdot,\cdot\rangle)$ *be a nonzero finite dimensional inner product space. Show that* $(H, \langle\cdot,\cdot\rangle)$ *is isometrically isomorphic to* $(\ell^2_n, \langle\cdot,\cdot\rangle_2)$*, where* $n := \dim(H)$*.*

More generally, one can prove that every nonzero Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is isometrically isomorphic to $(\ell^2(\Gamma), \langle\cdot,\cdot\rangle_2)$ for some nonempty set Γ.

Unlike the case of a general Banach space, one can give a satisfactory description of all the bounded linear functionals on a Hilbert space.

Theorem 3.20 (Riesz's Representation Theorem)**.** *Let* x^* *be a bounded linear functional on a Hilbert space* $(H, \langle\cdot,\cdot\rangle)$*. Then there exists an element* $x_0 \in H$ *such that* $x^*(y) = \langle y, x_0\rangle$ *for all* $y \in H$*. Moreover, the element* x_0 *is unique and the operator norm of* x^* *equals* $\|x_0\|$*.*

Proof. Consider the mapping $T : H \to H^*$ defined by, $T(x) := \langle\cdot, x\rangle$, i.e., $T(x)(y) = \langle y, x\rangle$ for each $y \in H$. From our earlier work we know that T well-defined, i.e., $T(x)$ is a continuous linear functional on H, for each $x \in H$. Fix $x \in H$, from the Cauchy-Schwarz inequality we have that $|\langle y, x\rangle| \leqslant \|x\|\|y\|$ for all $y \in H$ and so the operator norm of $T(x)$ is less than, or equal to, $\|x\|$. However, $|\langle x, x\rangle| = \|x\|^2$ and so

$$\|T(x)\| = \sup\{|T(x)(y)| : y \in S_H\} = \|x\|.$$

Thus, it remains to show that T is onto. To this end let $x^* \in H^* \setminus \{0\}$ and let $M := \mathrm{Ker}(x^*)$. Choose $x \in M^\perp \setminus \{0\}$. Note that this is possible since $M \neq H$. We claim that $H = \mathrm{span}\{x, M\}$, i.e., $H = \{h \in H : h = \lambda x + m \text{ for some } \lambda \in \mathbb{K} \text{ and } m \in M\}$. To prove this assertion, let us consider any $h \in H$. Then $h - [x^*(h)/x^*(x)]x \in M$ since

$$x^*(h - [x^*(h)/x^*(x)]x) = x^*(h) - [x^*(h)/x^*(x)]x^*(x) = x^*(h) - x^*(h) = 0.$$

Therefore, $h = [x^*(h)/x^*(x)]x + m$ where, $m := h - [x^*(h)/x^*(x)]x \in M$. We can now check that $T(x_0) = x^*$ where $x_0 := \mu x$ and $\mu := \overline{x^*(x)}/\|x\|^2$. But this is easy to check since we need only show that $T(x_0) = x^*$ on a spanning set for H. In particular, we need only show that $T(x_0) = x^*$ on $\{x\}$ and M. However, $T(x_0)(x) = \langle x, x_0\rangle = x^*(x)$ and $T(x_0)(m) = \langle m, x_0\rangle = 0 = x^*(m)$ for each $m \in M$. $\square$

For the idea behind this proof note that if f and g are linear functionals on a vector space V and $\mathrm{Ker}(f) \subseteq \mathrm{Ker}(g)$, then $g = \lambda f$ for some $\lambda \in \mathbb{K}$. We shall examine the structure of Hilbert spaces more closely later in this course.

Chapter 4

Hahn-Banach Theorem

A real-valued function p defined on a vector space V is called **sublinear** if for every $x, y \in V$ and $0 \leqslant \lambda < \infty$, $p(\lambda x) = \lambda p(x)$ and $p(x+y) \leqslant p(x) + p(y)$. If, moreover, $p(\lambda x) = |\lambda| p(x)$ for all $x \in V$ and all $\lambda \in \mathbb{K}$, then p is called a **semi-norm** on V.

Exercise 4.1. *(a) Show that every sublinear function p defined on a vector space V is convex and has the property that $p(0) = 0$.*

(b) Show that if p is a semi-norm then $p(x) = p(-x)$ for all $x \in V$ and $0 \leqslant p(x)$ for all $x \in V$. Hint : $0 = (1/2)(-x) + (1/2)x$.

Let us start this section with some linear algebra. Suppose that U is a subspace of a vector space $(V; +; \cdot\,)$, over the field of real numbers and suppose that $f : U \to \mathbb{R}$ is a linear mapping. We will look at possible "extensions" of f to larger subspaces of V. To this end, suppose that $x_0 \in V \setminus U$ and $W := \mathrm{span}(U, x_0)$. Then every $x \in W$ can be uniquely expressed in the form: $x = \lambda x_0 + u$ where $u \in U$ and $\lambda \in \mathbb{R}$. That is,

$$\mathrm{span}(U, x_0) = \{\lambda x_0 + u \in V : u \in U \text{ and } \lambda \in \mathbb{R}\}$$

and if, $\lambda_1 x_0 + u_1 = \lambda_2 x_0 + u_2$, then $\lambda_1 = \lambda_2$ and $u_1 = u_2$. For each $\alpha \in \mathbb{R}$, let $F_\alpha : W \to \mathbb{R}$ be defined by, $F_\alpha(x) := f(u) + \lambda\alpha$, where $x = \lambda x_0 + u$. Note that since the $\lambda \in \mathbb{R}$ and $u \in U$ are unique, this function is well-defined. It is also evident that $F_\alpha|_U = f$. It is also easy to verify that F_α is linear on W. Thus, each F_α is a linear extension of f to W.

Let us also observe that if $G : W \to \mathbb{R}$ is any linear function on W such that $G|_U = f$, then $G = F_\alpha$ for some $\alpha \in \mathbb{R}$. In fact, $G = F_{G(x_0)}$. To see this we simply do a calculation. Suppose that $G : W \to \mathbb{R}$ is a linear function such that $G|_U = f$ and $x \in W$. Then $x = \lambda x_0 + u$ for some unique $\lambda \in \mathbb{R}$ and $u \in U$ and

$$G(x) = G(\lambda x_0 + u) = \lambda G(x_0) + G(u) = \lambda G(x_0) + f(u) = F_{G(x_0)}(\lambda x_0 + u) = F_{G(x_0)}(x).$$

Next, we shall consider whether we can extend f to a linear function $G : W \to \mathbb{R}$ in such a way that if $f(u) \leqslant p(u)$ for all $u \in U$, then $G(x) \leqslant p(x)$ for $x \in W$, where $p : V \to \mathbb{R}$ is some sublinear functional on V. From our observations above this reduces to the question of whether there exist an $\alpha \in \mathbb{R}$ such that $F_\alpha(x) \leqslant p(x)$ for all $x \in W$, whenever, $f(u) \leqslant p(u)$ for all $u \in U$.

We shall look at this more closely. Firstly, $F_\alpha(x) \leqslant p(x)$ for all $x \in W$ if, and only if, $f(u) + \lambda\alpha \leqslant p(u + \lambda x_0)$ for all $u \in U$ and all $\lambda \in \mathbb{R}$ and this holds if, and only if,

$$\begin{aligned} f(u) + \lambda\alpha &\leqslant p(u + \lambda x_0) \text{ for all } u \in U \text{ and all } 0 \leqslant \lambda \text{ and} \\ f(u) + (-\lambda)\alpha &\leqslant p(u + (-\lambda)x_0) \text{ for all } u \in U \text{ and all } 0 < \lambda. \end{aligned}$$

Since $f(u) \leqslant p(u)$ for all $u \in U$, the above inequalities hold if, and only if,

$$\begin{aligned} \alpha &\leqslant p(\lambda^{-1}u + x_0) - f(\lambda^{-1}u) \text{ for all } u \in U \text{ and all } 0 < \lambda \text{ and} \\ \alpha &\geqslant f(\lambda^{-1}u) - p(\lambda^{-1}u - x_0) \text{ for all } u \in U \text{ and all } 0 < \lambda. \end{aligned}$$

Therefore, $F_\alpha(x) \leqslant p(x)$ for all $x \in W$, if, and only if,

$$\sup_{\substack{u\in U\\ 0<\lambda}} f(\lambda^{-1}u) - p(\lambda^{-1}u - x_0) \leqslant \alpha \leqslant \inf_{\substack{u\in U\\ 0<\lambda}} p(\lambda^{-1}u + x_0) - f(\lambda^{-1}u). \qquad (*)$$

Lemma 4.2. *Let U be a subspace of a vector space V over the real numbers and let $p : V \to \mathbb{R}$ be a sublinear functional on V. If f is a linear functional on U, $f(u) \leqslant p(u)$ for all $u \in U$ and $x_0 \in V \setminus U$, then there exists a linear function $G : \operatorname{span}(U, x_0) \to \mathbb{R}$ such that $G|_U = f$ and $G(x) \leqslant p(x)$ for all $x \in \operatorname{span}(U, x_0)$.*

Proof. Let $W := \operatorname{span}(U, x_0)$ and let $F_\alpha : W \to \mathbb{R}$ be defined by, $F_\alpha(x) := f(u) + \lambda\alpha$, where $u \in U$, $\lambda \in \mathbb{R}$ and $x = \lambda x_0 + u$. We need to show that the equality $(*)$ holds. Let $u_1, u_2 \in U$ and $\lambda_1, \lambda_2 \in (0, \infty)$. Then, since p is subadditive

$$f(\lambda_1^{-1}u_1 + \lambda_2^{-1}u_2) \leqslant p(\lambda_1^{-1}u_1 + \lambda_2^{-1}u_2) \leqslant p(\lambda_1^{-1}u_1 - x_0) + p(\lambda_2^{-1}u_2 + x_0).$$

Therefore,

$$f(\lambda_1^{-1}u_1) - p(\lambda_1^{-1}u_1 - x_0) \leqslant p(\lambda_2^{-1}u_2 + x_0) - f(\lambda_2^{-1}u_2).$$

Hold u_2 and λ_2 fixed, then

$$\sup_{\substack{u\in U\\ 0<\lambda}} f(\lambda^{-1}u) - p(\lambda^{-1}u - x_0) \leqslant p(\lambda_2^{-1}u_2 + x_0) - f(\lambda_2^{-1}u_2).$$

Now, take the infimum over $u_2 \in U$ and $0 < \lambda_2$ to get

$$\sup_{\substack{u\in U\\ 0<\lambda}} f(\lambda^{-1}u) - p(\lambda^{-1}u - x_0) \leqslant \inf_{\substack{u\in U\\ 0<\lambda}} p(\lambda^{-1}u + x_0) - f(\lambda^{-1}u).$$

Next, choose $\alpha \in \mathbb{R}$ such that

$$\sup_{\substack{u\in U\\ 0<\lambda}} f(\lambda^{-1}u) - p(\lambda^{-1}u - x_0) \leqslant \alpha \leqslant \inf_{\substack{u\in U\\ 0<\lambda}} p(\lambda^{-1}u + x_0) - f(\lambda^{-1}u).$$

Then, by $(*)$, $F_\alpha(x) \leqslant p(x)$ for all $x \in W$. $\square$

Theorem 4.3 (Hahn-Banach Theorem)**.** *Let U be a subspace of a vector space V (over $\mathbb{R}$) and let $p : V \to \mathbb{R}$ be a sublinear functional on V. If f is a linear functional on U and $f(u) \leqslant p(u)$ for all $u \in U$, then there exists a linear functional $F : V \to \mathbb{R}$ such that $F|_U = f$ and $F(x) \leqslant p(x)$ for all $x \in V$.*

Proof. Let $\mathscr{P}$ denote the collection of all ordered pairs (W', f'), where W' is a subspace of V containing U and $f' : W' \to \mathbb{R}$ is a linear functional defined on W' such that $f'|_U = f$ and satisfies $f'(x) \leqslant p(x)$ for all $x \in W'$. $\mathscr{P}$ is non-empty as $(U, f) \in \mathscr{P}$. We partially order $\mathscr{P}$ by, $(W', f') \leqslant (W'', f'')$ if $W' \subseteq W''$ and $f''|_{W'} = f'$. If $\{(W_\alpha, f_\alpha) : \alpha \in A\}$ is a nonempty totally ordered sub-family of $\mathscr{P}$, then $W' := \bigcup\{W_\alpha : \alpha \in A\}$ is a subspace of V containing U. The function $f' : W' \to \mathbb{R}$ defined by, $f'(x) := f_\alpha(x)$ if $x \in W_\alpha$ is well-defined and linear. In fact, $(W', f') \in \mathscr{P}$. Moreover, $(W_\alpha, f_\alpha) \leqslant (W', f')$ for all $\alpha \in A$. Therefore, by Zorn's Lemma, $\mathscr{P}$ has a maximal element (W, F). We must show that $W = V$. So suppose, in order to obtain a contradiction, that $W \neq V$ and pick $x_0 \in V \setminus W$. Then, by the previous lemma, there exists a linear function $G : \mathrm{span}(W, x_0) \to \mathbb{R}$ such that $G|_W = F$ and $G(x) \leqslant p(x)$ for all $x \in \mathrm{span}(W, x_0)$. Then $(\mathrm{span}(W, x_0), G) \in \mathscr{P}$ and so $(W, F) < (\mathrm{span}(W, x_0), G)$; which is impossible, since (W, F) is a maximal element of $\mathscr{P}$. Therefore, $W = V$, which completes the proof. $\square$

Exercise 4.4. *Let Y be a subspace of a normed linear space $(X, \|\cdot\|)$ (over $\mathbb{R}$). If $f \in Y^*$ then there exists an $F \in X^*$ such that $F|_Y = f$ and $\|F\| = \|f\|$. Hint: Consider $p : X \to \mathbb{R}$ defined by, $p(x) := \|f\| \cdot \|x\|$. Note also that $F(x) \leqslant p(x)$ for all $x \in X$ if, and only if, $|F(x)| \leqslant p(x)$ for all $x \in X$.*

Let V be a vector space over $\mathbb{C}$. Then V may also be considered as a vector space over $\mathbb{R}$ (or indeed, any subfield of $\mathbb{C}$). Let us denote this vector space by $V_\mathbb{R}$. In this way, if $(X, \|\cdot\|)$ is a normed linear space over $\mathbb{C}$ then $(X_\mathbb{R}, \|\cdot\|)$ is a normed linear space over $\mathbb{R}$. If $f \in X^*$ then $f_\mathbb{R} : X_\mathbb{R} \to \mathbb{R}$ defined by, $f_\mathbb{R}(x) := \mathrm{Real}[f(x)]$, is a member of $(X_\mathbb{R})^*$ (i.e., $f_\mathbb{R}$ is **real linear** and continuous).

Fact : Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{C}$ and let $f \in X^*$. Then $\|f_\mathbb{R}\| = \|f\|$. Clearly, $\|f_\mathbb{R}\| \leqslant \|f\|$. To obtain the reverse inequality, let us fix $x \in S_X$ and set $\theta := \arg(f(x)) \in [0, 2\pi)$. Then, $f(e^{-i\theta}x) = e^{-i\theta}f(x) \in \mathbb{R}$ and so $f(e^{-i\theta}x) = f_\mathbb{R}(e^{-i\theta}x)$. Therefore,

$$|f(x)| = |f(e^{-i\theta}x)| = |f_\mathbb{R}(e^{-i\theta}x)| \leqslant \|f_\mathbb{R}\|\|e^{-i\theta}x\| = \|f_\mathbb{R}\|\|x\| = \|f_\mathbb{R}\|.$$

Since $x \in S_X$ was arbitrary, $\|f\| = \sup_{x \in S_X} |f(x)| \leqslant \|f_\mathbb{R}\|$.

Exercise 4.5. *Let f be a linear functional defined on a vector space V over $\mathbb{C}$. Show that $f(x) = f_\mathbb{R}(x) - if_\mathbb{R}(ix)$ for all $x \in V$. Hint: Write f as: $f = f_\mathbb{R} + if_\mathbb{I}$ where $f_\mathbb{I}(x) := Im[f(x)]$ for all $x \in V$. Conversely, show that if g is a real linear functional on V and $f : V \to \mathbb{C}$ is defined by, $f(x) := g(x) - ig(ix)$ then f is* **complex linear** *and $f_\mathbb{R} = g$.*

Theorem 4.6. *Let Y be a subspace of a normed linear space $(X, \|\cdot\|)$ (over $\mathbb{C}$). If $f \in Y^*$ then there exists an $F \in X^*$ such that $F|_Y = f$ and $\|F\| = \|f\|$.*

Proof : Consider the real linear functional $f_{\mathbb{R}} : Y \to \mathbb{R}$. By an earlier exercise there exists a $G \in (X_{\mathbb{R}})^*$ such that $G|_Y = f_{\mathbb{R}}$ and $\|G\| = \|f_{\mathbb{R}}\| = \|f\|$. Define, $F : X \to \mathbb{C}$ by, $F(x) := G(x) - iG(ix)$. Then F is complex linear and

$$\|F\| = \|F_{\mathbb{R}}\| = \|G\| = \|f\|.$$

Moreover,

$$F|_Y(y) = G|_Y(y) - iG|_Y(iy) = f_{\mathbb{R}}(y) - if_{\mathbb{R}}(iy) = f(y)$$

for all $y \in Y$, i.e., $F|_Y = f$. $\square$

Corollary 4.7. *Let $(X, \|\cdot\|)$ be a normed linear space. For every $x \in X \setminus \{0\}$ there exists an $f \in S_{X^*}$ such that $f(x) = \|x\|$.*

Proof. Let $Y := \text{span}\{x\}$ and define $f \in Y^*$ by, $f(\lambda x) := \lambda\|x\|$. Clearly, $\|f\| = 1$ and $f(x) = \|x\|$. By Theorem 4.6 there exists an $F \in X^*$ such that $\|F\| = \|f\|$ and $F|_Y = f$. In particular, $F(x) = f(x) = \|x\|$. $\square$

Let S be a nonempty subset of a vector space V. We shall say that a point $x \in S$ is a **core point** of S if for every $v \in V$ there exists a $0 < \delta < \infty$ such that $x + \lambda v \in S$ for all $0 \leqslant \lambda < \delta$. The set of all core points of S is called the **core** of S.

Let C be a convex set in a vector space V with 0 in the core of C. Then the functional $\mu_C : V \to \mathbb{R}$ defined by, $\mu_C(x) := \inf\{\lambda > 0 : x \in \lambda C\}$ is called the **Minkowski functional** generated by the set C.

Theorem 4.8. *Let A be a convex subset of a vector space V with* 0 *in the core of A. Then $\mu_A : V \to \mathbb{R}$ is a sublinear functional. Moreover,*

$$\{x \in V : \mu_A(x) < 1\} \subseteq A \subseteq \{x \in V : \mu_A(x) \leqslant 1\}.$$

Proof. To show that μ_A is **positively homogeneous** (i.e., $\mu_A(sx) = s\mu_A(x)$ for all $0 \leqslant s < \infty$ and all $x \in V$) it is sufficient to show that $\mu_A(sx) \leqslant s\mu_A(x)$ for all $0 < s < \infty$ and all $x \in V$. To see this, let $0 < s < \infty$ and let $x \in V$, then

$$\mu_A(x) = \mu_A(s^{-1}(sx)) \leqslant s^{-1}\mu_A(sx) \quad \text{and so} \quad s\mu_A(x) \leqslant \mu_A(sx).$$

Note that as $\mu_A(0) = 0$, we get for free that $\mu_A(0x) = 0\mu_A(x)$ for all $x \in V$.

Next, let $0 < s < \infty$, $x \in V$ and let $0 < \varepsilon$. Then choose $0 < \lambda < (\mu_A(x) + \varepsilon/s)$ such that $x \in \lambda A$. Therefore, $sx \in (s\lambda)A$. Thus, $\mu_A(sx) \leqslant s\lambda$ and so $\mu_A(sx) \leqslant s\mu_A(x) + \varepsilon$. Since $0 < \varepsilon$ was arbitrary, $\mu_A(sx) \leqslant s\mu_A(x)$.

We now show that μ_A is **subadditive** (i.e., $\mu_A(x + y) \leqslant \mu_A(x) + \mu_A(y)$ for all $x, y \in V$). Let $x, y \in V$. Let $0 < \varepsilon$ be arbitrary. Then there exists $0 < \lambda_1 < \mu_A(x) + \varepsilon/2$ and $0 < \lambda_2 < \mu_A(y) + \varepsilon/2$ such that $x \in \lambda_1 A$ and $y \in \lambda_2 A$. Then

$$x + y \in \lambda_1 A + \lambda_2 A = (\lambda_1 + \lambda_2)A, \quad \text{since } A \text{ is convex.}$$

Therefore, $\mu_A(x+y) \leqslant \lambda_1 + \lambda_2 < \mu_A(x) + \mu_A(y) + \varepsilon$. Since $0 < \varepsilon$ was arbitrary, $\mu_A(x+y) \leqslant \mu_A(x) + \mu_A(y)$.

If $\mu_A(x) < 1$, then there exists a $0 < \lambda < 1$ such that $x \in \lambda A$ or $\lambda^{-1}x \in A$. Therefore, $x = \lambda[(\lambda^{-1}x)] + (1-\lambda)0 \in A$, since A is convex. On the other hand, if $x \in A$, then $x \in 1A$ and so $\mu_A(x) \leqslant 1$. $\square$

We now introduce some topology to the situation.

Proposition 4.9. *Let $p : X \to \mathbb{R}$ be a sublinear functional defined on a normed linear space $(X, \|\cdot\|)$. Then p is continuous on X if, and only if, p is continuous at 0.*

Proof. Clearly if p is continuous on X, then p is continuous at 0. So we consider the converse. Suppose that p is continuous at 0. Note that for any $x, y \in X$

$$p(x) \leqslant p(x-y) + p(y) \quad \text{and} \quad p(y) \leqslant p(y-x) + p(x).$$

Therefore, $p(x) - p(y) \leqslant p(x-y)$ and $p(y) - p(x) \leqslant p(y-x)$. Thus,

$$\pm[p(x) - p(y)] \leqslant \max\{p(x-y), p(y-x)\}.$$

That is, $|p(x) - p(y)| \leqslant \max\{p(x-y), p(y-x)\}$. Now, suppose $x_0 \in X$ and $0 < \varepsilon$ are given. Since p is continuous at 0, there exists a $0 < \delta$ such that $|p(x)| = |p(x) - p(0)| < \varepsilon$ for all $\|x\| = \|x - 0\| < \delta$. Note also that $|p(-x)| < \varepsilon$ for all $\|x\| < \delta$. So if, $\|x - x_0\| < \delta$, then

$$|p(x) - p(x_0)| \leqslant \max\{p(x-x_0), p(x_0-x)\} = \max\{p(x-x_0), p(-(x-x_0))\} < \varepsilon.$$

Hence, p is continuous at x_0. $\square$

Proposition 4.10. *Let A be a convex subset of a normed linear space $(X, \|\cdot\|)$. If $0 \in \mathrm{int}(A)$, then μ_A is continuous on X.*

Proof. By the Proposition 4.9 we need only show that μ_A is continuous at $0 \in X$. To this end, let $0 < \varepsilon$. Since $0 \in \mathrm{int}(A)$ there exists an $0 < r$ such that $rB_X \subseteq A$. Therefore, $(\varepsilon r)B_X \subseteq \varepsilon A$ and so if $x \in B(0, \varepsilon r)$, then $\mu_A(x) \leqslant \varepsilon$. Let $\delta := \varepsilon r$. Then $0 < \delta$ and if $\|x - 0\| < \delta$, $|\mu_A(x) - \mu_A(0)| = \mu_A(x) \leqslant \varepsilon$. This completes the proof. $\square$

Corollary 4.11. *Let A be a convex subset of a normed linear space $(X, \|\cdot\|)$ with $0 \in \mathrm{Cor}(A)$. Then μ_A is continuous on X if, and only if, $0 \in \mathrm{int}(A)$.*

Proof. From Proposition 4.10, if $0 \in \mathrm{int}(A)$, then μ_A is continuous on X. So we consider the converse. Suppose that μ_A is continuous on X, then $(\mu_A)^{-1}((-1,1))$ is an open subset of A and moreover, since $\mu_A(0) = 0 \in (-1,1)$, $(\mu_A)^{-1}((-1,1))$ is an open neighbourhood of 0, that is contained in A. Hence, $0 \in \mathrm{int}(A)$. $\square$

Exercise 4.12. *Let C be a convex subset of a normed linear space $(X, \|\cdot\|)$ with $0 \in int(C)$. Then*
$$\{x \in V : \mu_C(x) < 1\} = int(C) \quad and \quad \{x \in V : \mu_C(x) \leqslant 1\} = \overline{C}.$$
In particular, if C is a closed subset of X with $0 \in int(C)$ and $x_0 \notin C$, then $1 < \mu_C(x_0)$.

Theorem 4.13 (Separation Theorem))**.** *Let C be a nonempty closed convex subset of a normed linear space $(X, \|\cdot\|)$. If $x_0 \notin C$, then there exists an $f \in X^*$ such that* $\sup\{Real[f(x)] : x \in C\} < Real[f(x_0)]$.

Proof. First, let us consider the case when $(X, \|\cdot\|)$ is a normed linear space over $\mathbb{R}$. We may assume without loss of generality that $0 \in C$; otherwise we consider $C - x$ and $x_0 - x$ for some $x \in C$. Let $\delta := \text{dist}(x_0, C) > 0$. Set $D := \{x \in X : \text{dist}(x, C) \leqslant \delta/2\}$. Since $0 \in C$, we have that $0 \in \text{int}(D)$. D is also closed and convex and $x_0 \notin D$. Let μ_D be the Minkowski functional for D. Since D is closed and $x_0 \notin D$ we have $\mu_D(x_0) > 1$. Define a linear functional on $\text{span}\{x_0\}$ by, $f(\lambda x_0) := \lambda \mu_D(x_0)$. Then on $\text{span}\{x_0\}$ we have that $f(\lambda x_0) \leqslant \mu_D(\lambda x_0)$. Indeed, for $0 \leqslant \lambda$ it is clear from the definition of f; whereas for $\lambda < 0$ we have $f(\lambda x_0) = \lambda \mu_D(x_0) < 0$ while $\mu_D(\lambda x_0) \geqslant 0$. Extend f onto X so that $f(x) \leqslant \mu_D(x)$ for all $x \in X$. If $x \in D$, then $\mu_D(x) \leqslant 1$ and thus, $f(x) \leqslant \mu_D(x) \leqslant 1$. Since D contains a neighbourhood of the origin we have that f is a bounded on a neighbourhood of 0 and so $f \in X^*$. Since $f(x_0) = \mu_D(x_0) > 1$ we get that $\sup\{f(x) : x \in C\} \leqslant 1 < f(x_0)$.

In the complex case, we construct g from $(X_{\mathbb{R}})^*$ as in the real case and then define $f(x) := g(x) - ig(ix)$. ☐

Two subsets A and B of a normed linear space $(X, \|\cdot\|)$ are said to be **strongly separated by a closed hyperplane** if there exists an $f \in X^*$, an $\alpha \in \mathbb{R}$ and an $0 < \varepsilon < \infty$ such that:
$$A \subseteq \{x \in X : \text{Real}[f(x)] \leqslant \alpha - \varepsilon\} \quad \text{and} \quad B \subseteq \{x \in X : \text{Real}[f(x)] \geqslant \alpha + \varepsilon\}.$$

Theorem 4.14 (Strong Separation Theorem)**.** *Two disjoint closed and convex subsets A and B of a normed linear space $(X, \|\cdot\|)$ can be strongly separated by a closed hyperplane if there exists a $0 < \delta < \infty$ such that $(A + \delta B_X) \cap B = \varnothing$.*

Proof. Let $K := \overline{A - B}$. Then K is a nonempty closed and convex subset of X and $0 \notin K$. So from Theorem 4.13 there exists an $f \in X^*$ and an $r \in \mathbb{R}$ such that
$$\sup\{\text{Real}[f(x)] : x \in K\} < r < \text{Real}[f(0)] = 0.$$
In particular, for any $a \in A$ and $b \in B$, $\text{Real}[f(a) - f(b)] < r < 0$, or equivalently, $\text{Real}[f(a)] < r + \text{Real}[f(b)]$ for any $a \in A$ and $b \in B$. Hold $b \in B$ fixed and take the supremum over $a \in A$ to get:
$$\sup\{\text{Real}[f(a)] : a \in A\} \leqslant r + \text{Real}[f(b)]$$
Now take the infimum over $b \in B$ to get:
$$\sup\{\text{Real}[f(a)] : a \in A\} \leqslant r + \inf\{\text{Real}[f(b)] : b \in B\} < \inf\{\text{Real}[f(b)] : b \in B\}.$$

Let
$$\alpha := (1/2)\sup\{\text{Real}[f(a)] : a \in A\} + (1/2)\inf\{\text{Real}[f(b)] : b \in B\}$$
and $\varepsilon := (1/2)\inf\{\text{Real}[f(b)] : b \in B\} - (1/2)\sup\{\text{Real}[f(a)] : a \in A\} > 0$. Then,
$$A \subseteq \{x \in X : \text{Real}[f(x)] \leqslant \alpha - \varepsilon\} \quad \text{and} \quad B \subseteq \{x \in X : \text{Real}[f(x)] \geqslant \alpha + \varepsilon\}$$
which completes the proof. $\square$

Exercise 4.15. *Let M be a closed subspace of a normed linear space $(X, \|\cdot\|)$. If $x_0 \notin M$, then there exists an $f \in X^*$ such that Real$[f(x_0)] = 1$ and Real$[f(x)] = 0$ for all $x \in M$.*

Note that if a linear functional is bounded on a vector space, then it must be the zero functional. If the vector space is over the field of real number, and a linear function is bounded above (or below), then it must also be the zero functional.

Chapter 5

Baire's Theorem

Let C be a nonempty subset of a metric space (M, d). Then we define the **diameter of** C to be:

$$\operatorname{diam}(C) := \sup\{d(x, y) : x, y \in C\}.$$

Theorem 5.1 (Cantor Intersection Property). *Let $(F_n : n \in \mathbb{N})$ be a decreasing sequence (i.e., $F_{n+1} \subseteq F_n$ for all $n \in \mathbb{N}$) of nonempty closed subsets of a metric space (M, d). If (M, d) is a complete metric space and $\lim\limits_{n\to\infty} \operatorname{diam}(F_n) = 0$, then $\bigcap_{n\in\mathbb{N}} F_n \neq \varnothing$.*

Proof : For each $n \in \mathbb{N}$, choose $x_n \in F_n$. We claim that the sequence $(x_n : n \in \mathbb{N})$ is a Cauchy sequence. To verify this claim let us fix $0 < \varepsilon$. Since $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$, there exists an $N \in \mathbb{N}$ such that $\operatorname{diam}(F_n) < \varepsilon$ for all $n \geqslant N$. Let $N \leqslant n < m$, then

$$x_m \in F_m \subseteq F_{m-1} \subseteq \cdots \subseteq F_{n+1} \subseteq F_n \quad \text{i.e., } x_m, x_n \in F_n. \qquad (*)$$

Therefore, $d(x_m, x_n) \leqslant \operatorname{diam}(F_n) < \varepsilon$. This completes the proof of the claim.

Since (M, d) is a complete metric space, $(x_n : n \in \mathbb{N})$ converges to some point x_∞. We now claim that $x_\infty \in \bigcap_{n\in\mathbb{N}} F_n$. Let $n \in \mathbb{N}$, then by $(*)$ it follows that $x_m \in F_n$ for all $m \geqslant n$. Therefore, since F_n is a closed set, $x_\infty = \lim\limits_{m\to\infty} x_m \in F_n$. However, as $n \in \mathbb{N}$ was arbitrary, $x_\infty \in \bigcap_{n\in\mathbb{N}} F_n$. This completes the proof. $\square$

Theorem 5.2 (Baire Category Theorem). *Let (M, d) be a nonempty complete metric space and let $(O_n : n \in \mathbb{N})$ be dense open subsets of (M, d). Then, $\bigcap_{n=1}^\infty O_n$ is dense in (M, d).*

Proof. Let W be a nonempty open subset of (M, d); we will show that $\bigcap_{n=1}^\infty O_n \cap W \neq \varnothing$. We proceed inductively. First choose $x_1 \in O_1 \cap W$. Note this is possible since O_1 is dense in (M, d) and W is a nonempty open subset of (M, d). Then choose $0 < r_1 < 1$ such that $B[x_1, r_1] \subseteq O_1 \cap W$. Note: this is possible since $O_1 \cap W$ is an open set.

In general, we will choose $x_n \in M$ and $0 < r_n < 1/n$ such that

$$B[x_n, r_n] \subseteq B(x_{n-1}, r_{n-1}) \cap O_n \subseteq B[x_{n-1}, r_{n-1}].$$

Inductive step. Choose $x_{n+1} \in B(x_n, r_n) \cap O_{n+1}$. Note this is possible since O_{n+1} is dense in (M, d) and $B(x_n, r_n)$ is a nonempty open subset of (M, d). Then choose $0 < r_{n+1} < 1/(n+1)$ such that

$$B[x_{n+1}, r_{n+1}] \subseteq B(x_n, r_n) \cap O_{n+1} \subseteq B[x_n, r_n].$$

Note this is possible since $B(x_n, r_n) \cap O_{n+1}$ is open in (M, d).

By the Cantor Intersection Property $\varnothing \neq \bigcap_{n\in\mathbb{N}} B[x_n, r_n] \subseteq B[x_1, r_1] \subseteq W$. So we need only show that $\bigcap_{n\in\mathbb{N}} B[x_n, r_n] \subseteq \bigcap_{n\in\mathbb{N}} O_n$. However, by construction, $B[x_n, r_n] \subseteq O_n$ for all $n \in \mathbb{N}$, and so $\bigcap_{n\in\mathbb{N}} B[x_n, r_n] \subseteq \bigcap_{n\in\mathbb{N}} O_n$. This completes the proof. □

Example 5.3. *Let $\{r_n : n \in \mathbb{N}\}$ be an enumeration of the rational numbers. For each $n \in \mathbb{N}$, let $O_n := \mathbb{Q}\backslash\{r_n\}$. Then each O_n is a dense open subset of $\mathbb{Q}$. However, $\bigcap_{n=1}^{\infty} O_n = \varnothing$. This demonstrates the need for the metric space to be complete.*

Corollary 5.4. *Let (M, d) be a nonempty complete metric space and let $(F_n : n \in \mathbb{N})$ be a closed cover of (M, d). Then for some $k \in \mathbb{N}$, int$(F_k) \neq \varnothing$.*

Proof. For each $n \in \mathbb{N}$, let $O_n := M \setminus F_n$. Then,

$$\bigcap_{n\in\mathbb{N}} O_n = \bigcap_{n\in\mathbb{N}} (M \setminus F_n) = M \setminus \bigcup_{n\in\mathbb{N}} F_n = \varnothing.$$

Therefore, for some $k \in \mathbb{N}$, O_k must not be dense in (M, d). Thus, $F_k = M \setminus O_k$ has nonempty interior. □

We shall call a topological space (X, τ) a **Baire space** if for each sequence $(O_n : n \in \mathbb{N})$ of dense open subsets of (X, τ), $\bigcap_{n\in\mathbb{N}} O_n$ is dense in (X, τ).

From Theorem 5.2 we see that every nonempty complete metric space is a Baire space.

Exercise 5.5. *(a) Show that every nonempty regular compact space is a Baire space; (b) Show that if M is a nonempty complete metric space and X is a nonempty regular compact space, then $M \times X$ is a Baire space; (c) Show that every nonempty open subset of a Baire space is a Baire space (with the relative topology); (d) Show that if Y is a dense G_δ subset of a Baire space X, then Y (with the relative topology) is also a Baire space; (e) Let X be an uncountable set. Show that X with the co-finite (or co-countable) topology is a Baire space.*

Let (X, τ) be a topological space. Then we shall say that a subset F of (X, τ) is **first category** in (X, τ) if there exists a sequence $(F_n : n \in \mathbb{N})$ of closed subsets of (X, τ) such that: (i) $F \subseteq \bigcup_{n\in\mathbb{N}} F_n$ and (ii) $\text{int}(F_n) = \varnothing$ for each $n \in \mathbb{N}$. We shall say that a subset S of (X, τ) is **second category** if it is not first category.

Note that a topological space (X, τ) is a Baire space if, and only if, each nonempty open subset of (X, τ) is second category.

Lemma 5.6. *Suppose that $a < b$ and $f \in C[a,b]$. Then for each $\varepsilon > 0$ and $n \in \mathbb{N}$ there exists a piecewise linear mapping $g \in C[a,b]$ such that (i) $\|f - g\|_\infty < \varepsilon$ and (ii) $|g'_+(x)| > n$ for all $x \in [a,b)$.*

Proof. Consider the following set:

$$\mathcal{S} := \{x \in [a,b] : \text{ there exists a piecewise linear mapping } g \in C[a,x] \text{ with } g(x) = f(x), \|f|_{[a,x]} - g\|_\infty < \varepsilon \text{ and } |g'_+(y)| > n \text{ for all } y \in [a,x)\}$$

Let $s := \sup\{x \in [a,b] : x \in \mathcal{S}\}$. Clearly, $a < s \leqslant b$. To complete the proof we need to show that $s \in \mathcal{S}$ and that $s = b$ (i.e., show that $s < b$ leads to a contradiction). □

Example 5.7. *If $a < b$, then there exists a continuous nowhere differentiable function on $[a,b]$.*

Proof. Let $\mathcal{D}$ denote the set of all functions in $(C[a,b], \|\cdot\|_\infty)$ that have a right-hand derivative at some point of $[a,b)$. For each $n \in \mathbb{N}$, let

$$\mathcal{D}_n := \{f \in C[a,b] : \exists x \in [a, b - 1/n] \text{ for which } \sup_{0<h \leqslant 1/n} \left|\frac{f(x+h) - f(x)}{h}\right| \leqslant n\}.$$

Clearly, $\mathcal{D} \subseteq \bigcup_{n\in\mathbb{N}} \mathcal{D}_n$. Let us now show that each $\mathcal{D}_n$ is closed subset of $(C[a,b], \|\cdot\|_\infty)$. So fix $n \in \mathbb{N}$ and let $(f_k : k \in \mathbb{N})$ be a sequence in $\mathcal{D}_n$ that converges to f in $(C[a,b], \|\cdot\|_\infty)$. We need to show that $f \in \mathcal{D}_n$, i.e.,

$$\sup_{0<h \leqslant 1/n} \left|\frac{f(x+h) - f(x)}{h}\right| \leqslant n \quad \text{for some } x \in [a, b - 1/n].$$

Our first task is to find the candidate point $x \in [a, b-1/n]$ such that this inequality holds.

For each $i \in \mathbb{N}$, choose $x_i \in [a, b - 1/n]$ so that

$$\sup_{0<h \leqslant 1/n} \left|\frac{f_i(x_i+h) - f_i(x_i)}{h}\right| \leqslant n.$$

Since $[a, b - 1/n]$ is compact, by passing to a subsequence if needed, we may assume that $(x_i : i \in \mathbb{N})$ converges to $x \in [a, b - 1/n]$. (This is our candidate point!). We claim that:

$$\sup_{0<h \leqslant 1/n} \left|\frac{f(x+h) - f(x)}{h}\right| \leqslant n.$$

To see this, let $0 < h \leqslant 1/n$ be arbitrary. We need to show that

$$\left|\frac{f(x+h) - f(x)}{h}\right| \leqslant n.$$

To this end, let $0 < \varepsilon$ be arbitrary. We will show that:

$$\left|\frac{f(x+h) - f(x)}{h}\right| \leqslant n + \varepsilon.$$

Choose $k \in \mathbb{N}$ so that:

(i) $|f(x_k) - f(x)| < h\varepsilon/4$;

(ii) $|f(x_k + h) - f(x + h)| < h\varepsilon/4$ and

(iii) $\|f - f_k\| < h\varepsilon/4$.

Note that such a choice is possible since $x = \lim_{i\to\infty} x_i$ and $f = \lim_{i\to\infty} f_i$. Then,

$$\begin{aligned}\left|\frac{f(x+h)-f(x)}{h}\right| &\leqslant \left|\frac{f(x+h)-f(x_k+h)}{h}\right| + \left|\frac{f(x_k+h)-f_k(x_k+h)}{h}\right| \\ &+ \left|\frac{f_k(x_k+h)-f_k(x_k)}{h}\right| + \left|\frac{f_k(x_k)-f(x_k)}{h}\right| \\ &+ \left|\frac{f(x_k)-f(x)}{h}\right| \leqslant \varepsilon/4 + \varepsilon/4 + n + \varepsilon/4 + \varepsilon/4 = n + \varepsilon.\end{aligned}$$

Since $0 < \varepsilon$ was arbitrary,

$$\left|\frac{f(x+h)-f(x)}{h}\right| \leqslant n.$$

Since $h \in (0, 1/n]$ was arbitrary,

$$\sup_{0<h\leqslant 1/n} \left|\frac{f(x+h)-f(x)}{h}\right| \leqslant n.$$

This shows that $f \in \mathcal{D}_n$.

We now show that each $\mathcal{D}_n$ is nowhere dense in $(C[a,b], \|\cdot\|_\infty)$. So fix $n \in \mathbb{N}$. Suppose, in order to obtain a contradiction, that there is some $f \in \mathcal{D}_n$ and $r > 0$ such that $B(f,r) \subseteq \mathcal{D}_n$. Then, by Lemma 5.6, there exists a piecewise linear mapping $g : [a,b] \to \mathbb{R}$ such that (i) $\|f - g\|_\infty < r$; (ii) $g'_+(x)$ exists for all $x \in [a,b)$ and (iii) $|g'_+(x)| > n$ for all $x \in [a,b)$. However, this is impossible, since $g \in B(f,r) \subseteq \mathcal{D}_n$, but $g \notin \mathcal{D}_n$. $\square$

Remarks 5.8. *The previous example actually shows that the set of all functions in $C[a,b]$ that have a right-hand derivative at at-least one point of $[a,b)$ is of the first category in $(C[a,b], \|\cdot\|_\infty)$.*

Chapter 6

Open Mapping Theorem

Lemma 6.1. *Let* $(X, \|\cdot\|)$ *be a Banach space,* $(Y, |\!|\!|\cdot|\!|\!|)$ *a normed linear space and* $T \in B(X,Y)$*. If* $0 < r, s$ *satisfy* $B[0,s] \subseteq \overline{T(B[0,r])}$*, then* $B[0,s] \subseteq T(B[0,2r])$*.*

Proof. By considering the mapping $(r/s)T$ if necessary, we may assume that $r = s = 1$. Let y be an arbitrary element of $B[0,1]$. We will construct an $x \in B[0,2]$ such that $y = T(x)$.

Now, since $B[0,1] \subseteq \overline{T(B[0,1])}$, we have that for each $x \in X$ and each $0 < \varepsilon$

$$\begin{aligned} B(T(x),\varepsilon) &= T(x) + B(0,\varepsilon) \\ &= T(x) + \varepsilon B(0,1) \\ &\subseteq T(x) + \varepsilon\overline{T(B[0,1])} \\ &= \overline{T(x) + \varepsilon T(B[0,1])} = \overline{T(B[x,\varepsilon])}. \qquad (*) \end{aligned}$$

We shall inductively construct a sequence $(x_n : n \in \mathbb{N})$ in X such that:

(i) $x_n \in B[x_{n-1}, 1/2^{n-1}]$ for all $n \in \mathbb{N}$ and

(ii) $T(x_n) \in B(y, 1/2^n)$ for all $n \in \mathbb{N}$.

Set $x_0 := 0$. *Base Step.* Since $y \in \overline{T(B[0,1])}$, $B(y, 1/2^1) \cap T(B[0,1]) \neq \varnothing$. Choose $x_1 \in B[0,1] = B[x_0, 1/2^0]$ so that $T(x_1) \in B(y, 1/2^1)$.

Let $n \in \mathbb{N}$ and suppose that we have constructed $x_0, x_1, \ldots, x_n$ such that:

(i) $x_k \in B[x_{k-1}, 1/2^{k-1}]$ for all $1 \leqslant k \leqslant n$ and

(ii) $T(x_k) \in B(y, 1/2^k)$ for all $1 \leqslant k \leqslant n$.

Inductive Step. Since $T(x_n) \in B(y, 1/2^n)$, $y \in B(T(x_n), 1/2^n)$. Thus, by $(*)$, $y \in \overline{T(B[x_n, 1/2^n])}$. Therefore, $B(y, 1/2^{n+1}) \cap T(B[x_n, 1/2^n]) \neq \varnothing$. Hence, we may choose $x_{n+1} \in B[x_n, 1/2^n]$ such that $T(x_{n+1}) \in B(y, 1/2^{n+1})$. This completes the induction.

Now, $x_n = \sum_{k=1}^n (x_k - x_{k-1})$ and $\|x_k - x_{k-1}\| \leqslant 1/2^{k-1}$. Therefore,

$$x := \lim_{n\to\infty} x_n = \lim_{n\to\infty} \sum_{k=1}^n (x_k - x_{k-1}) \quad \text{exists and moreover,}$$

$$
\begin{aligned}
\|x\| = \|\lim_{n\to\infty} \textstyle\sum_{k=1}^{n}(x_k - x_{k-1})\| &= \lim_{n\to\infty} \|\textstyle\sum_{k=1}^{n}(x_k - x_{k-1})\| \\
&\leqslant \lim_{n\to\infty} \textstyle\sum_{k=1}^{n}\|x_k - x_{k-1}\| \\
&\leqslant \lim_{n\to\infty} \textstyle\sum_{k=1}^{n} 1/2^{k-1} = 2.
\end{aligned}
$$

i.e., $x \in 2B_X$. On the other hand, $|||y - T(x_n)||| \leqslant 1/2^n$ for all $n \in \mathbb{N}$. Therefore,

$$
\begin{aligned}
0 \leqslant |||y - T(x)||| &= |||y - T(\lim_{n\to\infty} x_n)||| \\
&= |||y - \lim_{n\to\infty} T(x_n)||| \\
&= \lim_{n\to\infty} |||y - T(x_n)||| \\
&\leqslant \lim_{n\to\infty} 1/2^n = 0.
\end{aligned}
$$

Thus, $y = T(x)$ and so $B[0,1] \subseteq T(B[0,2])$. $\square$

Theorem 6.2 (Open Mapping Theorem)**.** *Suppose that $(X, \|\cdot\|)$ and $(Y, |||\cdot|||)$ are Banach spaces and $T \in B(X,Y)$. If T maps onto Y, then T is an open mapping (i.e., maps open sets to open sets).*

Proof. First, let us show that there exists an $0 < s$ such that $B(0,s) \subseteq T(2B_X)$. In light of Lemma 6.1, to accomplish this, we need only show that $B(0,s) \subseteq \overline{T(B_X)}$. To this end, consider the following:

$$
Y = T(X) = T(\bigcup_{n\in\mathbb{N}} nB_X) = \bigcup_{n\in\mathbb{N}} nT(B_X) \subseteq \bigcup_{n\in\mathbb{N}} n\overline{T(B_X)} \subseteq Y.
$$

Therefore, by Baire's theorem, for some $n_0 \in \mathbb{N}$, $\mathrm{int}[n_0\overline{T(B_X)}] \neq \varnothing$. Choose $y \in Y$ and $r > 0$ such that $B[y,r] \subseteq n_0\overline{T(B_X)}$. Then,

$$
B[0,r] = (1/2)B[-y,r] + (1/2)B[y,r] \subseteq n_0\overline{T(B_X)}
$$

since $n_0\overline{T(B_X)}$ is convex and symmetric. Therefore, if $s := r/n_0$, then

$$
sB_Y = (1/n_0)B[0,r] \subseteq (1/n_0)(n_0\overline{T(B_X)}) = \overline{T(B_X)}.
$$

Next, let G be a nonempty open subset of $(X, \|\cdot\|)$ and let $y \in T(G)$. Choose $x \in G$ such that $y = T(x)$. Since G is open there exists a $\delta > 0$ such that $B[x, 2\delta] \subseteq G$. Then,

$$
\begin{aligned}
y \in B(y, s\delta) = y + \delta B(0,s) = T(x) + \delta B(0,s) &\subseteq T(x) + \delta T(2B_X) = T(x + 2\delta B_X) \\
&= T(B[x, 2\delta]) \subseteq T(G)
\end{aligned}
$$

and so $T(G)$ is open in $(Y, |||\cdot|||)$. $\square$

Corollary 6.3. *Suppose that $(X, \|\cdot\|)$ and $(Y, |||\cdot|||)$ are Banach spaces and $T \in B(X,Y)$. If T is 1-to-1 and onto, then $T^{-1} \in B(Y,X)$.*

Proof. Since T is 1-to-1 and onto T^{-1} exists and is linear. So it is sufficient to show that T^{-1} is continuous. To this end, let G be an open subset of $(X, \|\cdot\|)$. Then $(T^{-1})^{-1}(G) = T(G)$; which is open in $(Y, |||\cdot|||)$. Therefore, T^{-1} is continuous. $\square$

Corollary 6.4. *Suppose that $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$ are Banach spaces and $T \in B(X,Y)$. If T is onto, then $(X/Ker(T), \|\cdot\|)$ is isomorphic to $(Y, \|\!|\cdot|\!\|)$.*

Proof. Apply Corollary 6.3 to the mapping $\widehat{T} : X/\mathrm{Ker}(T) \to Y$ defined by, $\widehat{T}(x + \mathrm{Ker}(T)) := T(x)$. To see that $\widehat{T}$ is continuous, notice that the open unit ball in $X/\mathrm{Ker}(T)$ is contained in $\widehat{B_X}$ and so $\widehat{T}(B(0,1)) \subseteq T(B_X)$; which is bounded in $(Y, \|\!|\cdot|\!\|)$. □

Theorem 6.5 (Closed Graph Theorem)**.** *Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces and T is a linear mapping from X into Y. If the graph of T is a closed subset of $X \times Y$, then T is continuous.*

Proof. Let $\|\cdot\| : X \to \mathbb{R}$ be defined by, $\|x\| := \|x\|_X + \|T(x)\|_Y$. Then $\|\cdot\|$ is a norm on X and $\|x\|_X \leqslant \|x\|$ for all $x \in X$. Therefore, the linear mapping $I : (X, \|\cdot\|) \to (X, \|\cdot\|_X)$ defined by, $I(x) := x$ is 1-to-1, onto and continuous. Next, we will show that $(X, \|\cdot\|)$ is a Banach space. Now, if $(x_n : n \in \mathbb{N})$ is a Cauchy sequence in $(X, \|\cdot\|)$, then $(x_n : n \in \mathbb{N})$ is a Cauchy sequence in $(X, \|\cdot\|_X)$ and $(T(x_n) : n \in \mathbb{N})$ is a Cauchy sequence in $(Y, \|\cdot\|_Y)$. Let $x := \lim_{n\to\infty} x_n$ and $y := \lim_{n\to\infty} T(x_n)$. Since T has closed graph $y = T(x)$ (i.e., $(x,y) \in \mathrm{Graph}(T)$). Therefore,

$$\lim_{n\to\infty} \|x - x_n\| = \lim_{n\to\infty} \|x - x_n\|_X + \lim_{n\to\infty} \|T(x) - T(x_n)\|_Y = 0$$

and so $(X, \|\cdot\|)$ is a Banach space. Thus, by Corollary 6.3 $\|\cdot\|$ and $\|\cdot\|_X$ are equivalent norms on X; which implies that T is continuous. □

Exercise 6.6. *Show that if $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$ are normed linear spaces and $T \in B(X,Y)$, then T has a closed graph.*

Application

Let $(X, \|\cdot\|)$ be an infinite dimensional separable normed linear space. A sequence $(e_n : n \in \mathbb{N})$ in $(X, \|\cdot\|)$ is called a **Schauder basis** if for every $x \in X$ there exist *unique* scalars $(a_n : n \in \mathbb{N})$, called the **coordinates** of x, such that $x = \sum_{n\in\mathbb{N}} a_n e_n$. For each $n \in \mathbb{N}$, the **canonical projections** $P_n : X \to X$ are defined by, $P_n(\sum_{k=1}^{\infty} a_k e_k) = \sum_{k=1}^{n} a_k e_k$.

If $(e_k : k \in \mathbb{N})$ is a Schauder basis for a normed linear space $(X, \|\cdot\|)$, then for each $n \in \mathbb{N}$, the mapping $x_n^* : X \to \mathbb{K}$ defined by, $x_n^*(x) := a_n$, where a_n is the n^{th}-coordinate of x with respect to the basis $(e_k : k \in \mathbb{N})$, is a linear functional on $(X, \|\cdot\|)$, called the **coordinate functional**.

Theorem 6.7. *If $(e_k : k \in \mathbb{N})$ is a Schauder basis for a Banach space $(X, \|\cdot\|)$. Then for each $n \in \mathbb{N}$, the coordinate functional x_n^* is continuous.*

Proof. Define $\|\cdot\|_X : X \to \mathbb{R}$ by, $\|x\|_X := \sup\{\|P_n(x)\| : n \in \mathbb{N}\}$. (Note: this is well defined since $(P_n(x) : n \in \mathbb{N})$ converges to x in $(X, \|\cdot\|)$ and so $\sup\{\|P_n(x)\| : n \in \mathbb{N}\} < \infty$ i.e., convergent sequences are bounded.)

Then $\|\cdot\|_X$ is a norm on X. Moreover, $\|x\| \leqslant \|x\|_X$ for all $x \in X$ since

$$\|x\| = \|\lim_{n\to\infty} P_n(x)\| = \lim_{n\to\infty} \|P_n(x)\| \leqslant \sup\{\|P_n(x)\| : n \in \mathbb{N}\} = \|x\|_X.$$

Therefore, if we can show that $\|\cdot\|_X$ is a complete norm, then we have by Corollary 6.3 that $\|\cdot\|_X$ is an equivalent norm to $\|\cdot\|$. To show this we need several facts: (i) If $(x_n : n \in \mathbb{N})$ is a Cauchy sequence in $(X, \|\cdot\|_X)$, then for each $k \in \mathbb{N}$, $(x_k^*(x_n) : n \in \mathbb{N})$ is a Cauchy sequence in $\mathbb{K}$, and hence is convergent; (ii) If $a_k := \lim_{n\to\infty} x_k^*(x_n)$ for each $k \in \mathbb{N}$, then $x := \sum_{k\in\mathbb{N}} a_k e_k$ is an element of X; (iii) $(x_n : n \in \mathbb{N})$ converges to x in $(X, \|\cdot\|_X)$.

Since each x_n^* is continuous with respect to $\|\cdot\|_X$ and $\|\cdot\|_X$ is equivalent to $\|\cdot\|$ we have that each x_n^* is continuous with respect to $\|\cdot\|$. □

Exercise 6.8. *For each $n \in \mathbb{N}$, let $f_n : [0,1] \to \mathbb{R}$ be defined by, $f_n(x) := x^n$.*

(a) *Show that $(f_n : n \in \mathbb{N})$ is a Schauder basis for $(P[0,1], \|\cdot\|_\infty)$, i.e., the polynomials on $[0,1]$ equipped with the sup-norm.*

(b) *Show that the coordinate functionals on $P[0,1]$, with respect to the basis $(f_n : n \in \mathbb{N})$, are not continuous.*

Exercise 6.8 shows that the completeness of $(X, \|\cdot\|)$ is essential to deduce the continuity of the the coordinate functionals.

Chapter 7

Uniform Boundedness Theorem

Theorem 7.1 (Uniform Boundedness Theorem)**.** *Let* $(X, \|\cdot\|)$ *be a Banach space,* $(Y, \|\!|\cdot|\!\|)$ *be a normed linear space and* $\{T_\alpha : \alpha \in A\} \subseteq B(X, Y)$*. If*

$$\{x \in X : \{T_\alpha(x) : \alpha \in A\} \textit{ is bounded}\}$$

is second category in $(X, \|\cdot\|)$*, then* $\{T_\alpha : \alpha \in A\}$ *is uniformly bounded (i.e., there exists an* $M > 0$ *such that* $\|T_\alpha\| \leqslant M$ *for all* $\alpha \in A$*).*

Proof : Let $S := \{x \in X : \{T_\alpha(x) : \alpha \in A\} \text{ is bounded}\}$. For each $n \in \mathbb{N}$, let

$$\begin{aligned} F_n &:= \{x \in X : \|\!|T_\alpha(x)|\!\| \leqslant n \text{ for all } \alpha \in A\} \\ &= \bigcap_{\alpha \in A} \{x \in X : \|\!|T_\alpha(x)|\!\| \leqslant n\} \\ &= \bigcap_{\alpha \in A} (\|\cdot\| \circ T_\alpha)^{-1}([0, n]); \end{aligned}$$

which is closed. Since $\{T_\alpha : \alpha \in A\}$ is pointwise bounded on S, $S \subseteq \bigcup_{n=1}^{\infty} F_n$. Therefore, for some $n_0 \in \mathbb{N}$, $\operatorname{int}(F_{n_0}) \neq \varnothing$. Choose $x \in X$ and $r > 0$ such that $B[x, r] \subseteq F_{n_0}$. Then $B[-x, r] \subseteq F_{n_0}$ and $B[0, r] = \frac{1}{2}B[-x, r] + \frac{1}{2}B[x, r] \subseteq F_{n_0}$, since F_{n_0} is symmetric and convex. Hence, for any $x \neq 0$ and $\alpha \in A$

$$\frac{r}{\|x\|}\|\!|T_\alpha(x)|\!\| = \|\!|T_\alpha\left(\frac{xr}{\|x\|}\right)|\!\| \leqslant n_0.$$

Therefore, $\|\!|T_\alpha(x)|\!\| \leqslant (n_0/r)\|x\|$ for all $x \in X$ and $\alpha \in A$ and so $\|T_\alpha\| \leqslant M$ for all $\alpha \in A$, where $M := (n_0/r)$. $\square$

Corollary 7.2. *Let* $(X, \|\cdot\|)$ *be a Banach space,* $(Y, \|\!|\cdot|\!\|)$ *be a normed linear space and* $\{T_\alpha : \alpha \in A\} \subseteq B(X, Y)$*. If for some* $x_0 \in X$*,* $\{T_\alpha(x_0) : \alpha \in A\}$ *is unbounded, then* $\{x \in X : \{T_\alpha(x) : \alpha \in A\} \textit{ is bounded}\}$ *is first category in* $(X, \|\cdot\|)$*.*

Theorem 7.3 (Banach-Steinhaus Theorem)**.** *Let* $(X, \|\cdot\|)$ *and* $(Y, \|\!|\cdot|\!\|)$ *be Banach spaces and let* $(T_n : n \in \mathbb{N})$ *be a sequence in* $B(X, Y)$*. If* $(T_n : n \in \mathbb{N})$ *is pointwise Cauchy, then it is pointwise convergent to some* $T \in B(X, Y)$*.*

Proof. For each $x \in X$, let $T(x) := \lim_{n\to\infty} T_n(x)$. Since $(Y, \|\!| \cdot |\!\|)$ is complete, this is well-defined. Moreover, it is easy to check that T is linear. Since $(T_n : n \in \mathbb{N})$ is pointwise convergent it is pointwise bounded. Thus, by the Uniform Boundedness Theorem, there exists an $M > 0$ such that $\|T_n\| \leqslant M$ for all $n \in \mathbb{N}$. In particular, $\|T_n(x)\| \leqslant M\|x\|$ for all $x \in X$ and all $n \in \mathbb{N}$. Therefore, $\|T(x)\| \leqslant M\|x\|$ for all $x \in X$. □

Let $(X, \|\cdot\|)$ be a normed linear space. For each $x \in X$ we define, $\widehat{x} \in X^{**} := (X^*)^*$ by, $\widehat{x}(x^*) := x^*(x)$ for each $x^* \in X^*$. To show that $\widehat{x}$ is really in X^{**} we must first check that it is linear and then check that it is continuous. Fix $x \in X$ and suppose that x^* and y^* are in X^*, then

$$\widehat{x}(x^* + y^*) = (x^* + y^*)(x) = x^*(x) + y^*(x) = \widehat{x}(x^*) + \widehat{x}(y^*).$$

Also, if $s \in \mathbb{K}$ and $x^* \in X^*$, then we have that

$$\widehat{x}(sx^*) = (sx^*)(x) = sx^*(x) = s\widehat{x}(x^*).$$

This shows that $\widehat{x}$ is linear. Now, let $x^* \in X^*$, then $|\widehat{x}(x^*)| = |(x^*)(x)| \leqslant \|x^*\|\|x\|$. Therefore, $\|\widehat{x}\| \leqslant \|x\|$.

Proposition 7.4. *Let $(X, \|\cdot\|)$ be a normed linear space, then for each $x \in X$, $\|\widehat{x}\| = \|x\|$.*

Proof. Fix $x \in X$, then by Corollary 4.7, there existence of a continuous linear function $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \|x\|$. Therefore,

$$\|\widehat{x}\| \geqslant \frac{|\widehat{x}(x^*)|}{\|x^*\|} = |\widehat{x}(x^*)| = |x^*(x)| = \|x\|.$$

This completes the proof. □

Moreover, the mapping $x \mapsto \widehat{x}$ from X into X^{**} is linear. To see this, fix $x^* \in X^*$. Then,

$$\widehat{(x+y)}(x^*) = x^*(x+y) = x^*(x) + x^*(y) = \widehat{x}(x^*) + \widehat{y}(x^*).$$

This shows that $\widehat{x+y} = \widehat{x} + \widehat{y}$. Also, if $s \in \mathbb{K}$ and $x^* \in X^*$, then

$$\widehat{(sx)}(x^*) = x^*(sx) = sx^*(x) = s\widehat{x}(x^*),$$

which shows that $\widehat{(sx)} = s\widehat{x}$.

If $(X, \|\cdot\|)$ is a Banach space, then $\widehat{X}$ is a closed subspace of $(X^{**}, \|\cdot\|)$, where $\widehat{X}$ is defined as $\{\widehat{x} : x \in X\}$. We call $\widehat{X}$ the **natural embedding** of X into X^{**} and we call $x \mapsto \widehat{x}$ from X into X^{**} the **natural embedding mapping**.

We will say that a subset A of a normed linear space $(X, \|\cdot\|)$ is **weakly bounded** if for each $x^* \in X^*$, $\sup_{x\in A} |x^*(x)| < \infty$.

Theorem 7.5. *Let A be a nonempty subset of a normed linear space $(X, \|\cdot\|)$. Then A is a weakly bounded if, and only if, A is bounded.*

Proof. Suppose A is bounded (i.e., there exists an $M > 0$ such that $\|x\| \leqslant M$ for all $x \in A$). Then, $|x^*(x)| \leqslant \|x^*\| \cdot \|x\| \leqslant M\|x^*\| < \infty$ for all $x \in A$.

Conversely, suppose A is weakly bounded and consider the family, $\{\widehat{x} \in X^{**} : x \in A\}$. Now, $(X^*, \|\cdot\|)$ is a Banach space and by the hypothesis $\{\widehat{x} \in X^{**} : x \in A\}$ is pointwise bounded. Therefore, by the Uniform Boundedness Theorem, there exists an $M > 0$ such that $\|x\| = \|\widehat{x}\| \leqslant M$ for all $x \in A$. $\square$

Corollary 7.6. *Let T be a linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. Then T is continuous if, and only if, for each $y^* \in Y^*$, $y^* \circ T : X \to \mathbb{K}$ is continuous.*

Proof. If T is continuous, then for every $y^* \in Y^*$, $y^* \circ T$ is continuous. This follows from the general fact that the composition of continuous functions is continuous. Now, suppose that for each $y^* \in Y^*$, $y^* \circ T : X \to \mathbb{K}$ is continuous. We will show that $T(B_X)$ is a weakly bounded subset of $(Y, \|\!|\cdot|\!\|)$, and hence by Theorem 7.5, a bounded subset of $(Y, \|\!|\cdot|\!\|)$. Let $y^* \in Y^*$. Then $y^*(T(B_X)) = (y^* \circ T)(B_X)$ is a bounded subset of $\mathbb{K}$, since by assumption, $y^* \circ T$ is a bounded operator. Since $y^* \in Y^*$ was arbitrary, it follows that $T(B_X)$ is weakly bounded. $\square$

In the proof of the next theorem we will, in order to avoid any possible confusion, denote the norm on the second dual of the normed linear space $(X, \|\cdot\|)$ by $\|\cdot\|^{**}$.

Theorem 7.7. *Let $(X, \|\cdot\|)$ be a normed linear space. Then there exists a Banach space $(Y, \|\!|\cdot|\!\|)$ (called the* **completion** *of X) such that $(X, \|\cdot\|)$ is isometrically isomorphic to a dense subspace of $(Y, \|\!|\cdot|\!\|)$.*

Proof. Firstly, $(X^{**}, \|\cdot\|^{**})$ is a Banach space. Let $Y := \overline{\widehat{X}}$ and let $\|\!|\cdot|\!\|$ denote the restriction of the norm $\|\cdot\|^{**}$ to the subspace Y. Then, $(Y, \|\!|\cdot|\!\|)$ is a Banach space, $\widehat{X}$ is clearly dense in $(Y, \|\!|\cdot|\!\|)$ and $(X, \|\cdot\|)$ is isometrically isomorphic to $\widehat{X}$. $\square$

Application

Let $C_{2\pi}(\mathbb{R})$ denote the space of all continuous real-valued functions defined on $\mathbb{R}$ such that $f(x) = f(x + 2\pi)$ for all $x \in \mathbb{R}$. Note that it follows from induction that if $f \in C_{2\pi}(\mathbb{R})$, $x \in \mathbb{R}$ and $n \in \mathbb{Z}$, then $f(x) = f(x + 2\pi n)$.

It follow from this that for any $a < b$ and any $n \in \mathbb{Z}$, $\displaystyle\int_a^b f(t)\ \mathrm{d}t = \int_{a+2n\pi}^{b+2n\pi} f(t)\ \mathrm{d}t \quad (*)$.

Furthermore, if $0 \leqslant x < 2\pi$, then $\displaystyle\int_{-\pi}^{\pi} f(t)\ \mathrm{d}t = \int_{-\pi+x}^{\pi+x} f(t)\ \mathrm{d}t$. To see this consider the following.

$$\begin{aligned}
\int_{-\pi}^{\pi} f(t)\ \mathrm{d}t &= \int_{-\pi}^{-\pi+x} f(t)\ \mathrm{d}t + \int_{-\pi+x}^{\pi} f(t)\ \mathrm{d}t \\
&= \int_{\pi}^{\pi+x} f(t)\ \mathrm{d}t + \int_{-\pi+x}^{\pi} f(t)\ \mathrm{d}t \quad \text{apply } (*) \text{ with } a=-\pi \text{ and } b=-\pi+x \\
&= \int_{-\pi+x}^{\pi+x} f(t)\ \mathrm{d}t. \quad (**)
\end{aligned}$$

By combining $(*)$ and $(**)$ we get the (probably obvious) fact that for any $x \in \mathbb{R}$,

$$\int_{-\pi}^{\pi} f(t)\ \mathrm{d}t = \int_{-\pi+x}^{\pi+x} f(t)\ \mathrm{d}t.$$

Theorem 7.8. *There exists a function $f \in C_{2\pi}(\mathbb{R})$ whose Fourier series is divergent at each point of a dense subset of $\mathbb{R}$.*

Proof. We shall begin by showing that

$$\{f \in C_{2\pi}(\mathbb{R}) : \text{ the Fourier series for } f \text{ converges at } 0\}$$

is first category in $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$. For each $f \in C_{2\pi}(\mathbb{R})$ and $n \in \mathbb{N}$, the n^{th}-partial sum of the Fourier series of f is:

$$S_n(f,x) := \sum_{k=-n}^{n} c_k e^{ikx} \quad \text{where} \quad c_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\ \mathrm{d}t.$$

Now,

$$S_n(f,x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_n(t)\ \mathrm{d}t \quad \text{where,} \quad D_n(t) := \frac{\sin[(n+1/2)t]}{\sin[(1/2)t]}.$$

Notice that if we define $\varphi_n : C_{2\pi}(\mathbb{R}) \to \mathbb{R}$ by, $\varphi_n(f) := S_n(f,0)$ for each $n \in \mathbb{N}$, then each φ_n is a continuous linear functional on $C_{2\pi}(\mathbb{R})$. In fact,

$$\|\varphi_n\| = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\ \mathrm{d}t.$$

Next suppose, in order to obtain a contradiction, that

$$S := \{f \in C_0[-\pi,\pi] : \text{ the Fourier series for } f \text{ converges at } 0\}$$

is second category in $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$. Then by the Uniform Boundedness Theorem the set $\{\|\varphi_n\| : n \in \mathbb{N}\}$ is bounded. However, we have that,

$$
\begin{aligned}
\int_{-\pi}^{\pi} |D_n(t)|\ \mathrm{d}t &= 2\int_0^{\pi} |D_n(t)|\ \mathrm{d}t \geqslant 4\int_0^{\pi} \frac{|\sin[(n+1/2)t]|}{t}\ \mathrm{d}t \\
&= 4\int_0^{\pi/2} \frac{|\sin[(2n+1)t]|}{t}\ \mathrm{d}t \\
&\geqslant 4\sum_{k=0}^{n-1}\int_{\frac{k\pi}{2n+1}}^{\frac{(k+1)\pi}{2n+1}} \frac{|\sin[(2n+1)t]|}{t}\ \mathrm{d}t \\
&\geqslant 4\sum_{k=0}^{n-1}\frac{2n+1}{(k+1)\pi}\int_{\frac{k\pi}{2n+1}}^{\frac{(k+1)\pi}{2n+1}} |\sin[(2n+1)t]|\ \mathrm{d}t = \frac{8}{\pi}\sum_{k=0}^{n-1}\frac{1}{k+1};
\end{aligned}
$$

which is divergent. But this contradicts the boundedness of $\{\|\varphi_n\| : n \in \mathbb{N}\}$. So the set S must be first category in $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$.

Next, we show that for each $\alpha \in \mathbb{R}$,

$$S_\alpha := \{f \in C_{2\pi}(\mathbb{R}) : \text{ the Fourier series for } f \text{ converges at } -\alpha\}$$

is of the first category in $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$. To this end, fix $\alpha \in \mathbb{R}$. Let $T_\alpha : \mathbb{R} \to \mathbb{R}$ be defined by, $T_\alpha(t) := t + \alpha$ for each $t \in \mathbb{R}$ and let $T_\alpha^* : C_{2\pi}(\mathbb{R}) \to C_{2\pi}(\mathbb{R})$ be defined by, $T_\alpha^*(f) := f \circ T_\alpha$. Then T_α^* is an isometry. Hence, $T_\alpha^*(S)$ is first category in $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$.

Claim: $S_\alpha \subseteq T_\alpha^*(S)$. To see this consider $g \in S_\alpha$. Since T_α^* is onto there exists an $f \in C_{2\pi}(\mathbb{R})$ such that $g = T_\alpha^*(f)$. We need to show that $f \in S$. To this end, consider the following:

$$
\begin{aligned}
c_k &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-ikt}\ \mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} T_\alpha^*(f)(t)e^{-ikt}\ \mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t+\alpha)e^{-ikt}\ \mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi+\alpha}^{\pi+\alpha} f(t)e^{-ik(t-\alpha)}\ \mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ik(t-\alpha)}\ \mathrm{d}t = \frac{e^{ik\alpha}}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\ \mathrm{d}t.
\end{aligned}
$$

Therefore, for each $n \in \mathbb{N}$, we have that

$$
\begin{aligned}
S_n(g, -\alpha) &= \sum_{k=-n}^{n}\left(\frac{e^{ik\alpha}}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\ \mathrm{d}t\right)\cdot e^{-ik\alpha} \\
&= \sum_{k=-n}^{n}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\ \mathrm{d}t\right)\cdot e^{ik0} = S_n(f, 0).
\end{aligned}
$$

Therefore, $f \in S$.

Let $\mathcal{S} := \bigcup\{S_\alpha : \alpha \in \mathbb{Q}\}$. Then $\mathcal{S}$ is first category in $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$ and $\lim_{n\to\infty} S_n(f, \alpha)$ diverges for each $\alpha \in \mathbb{Q}$ and each $f \in C_{2\pi}(\mathbb{R}) \setminus \mathcal{S}$. $\square$

Exercise 7.9. *Let $n \in \mathbb{N}$ and $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$. Show that:*

$$\sum_{k=-n}^{n} e^{ikt} = \frac{e^{i(n+1)t} - e^{-int}}{e^{it} - 1} = \frac{e^{i(n+1/2)t} - e^{-i(n+1/2)t}}{e^{it/2} - e^{-it/2}} = \frac{\sin[(n+1/2)t]}{\sin[(1/2)t]}.$$

Remarks 7.10. *For each $f \in C_{2\pi}(\mathbb{R})$ and $n \in \mathbb{N}$,*

$$\begin{aligned}
S_n(f,x) &= \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} \, \mathrm{d}t \right) \cdot e^{ikx} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{k=-n}^{n} e^{ik(x-t)} \right) \mathrm{d}t \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_n(x-t) \, \mathrm{d}t, \quad \textit{where } D_n(t) := \frac{\sin[(n+1/2)t]}{\sin[(1/2)t]} \\
&= -\frac{1}{2\pi} \int_{x+\pi}^{x-\pi} f(x-t') D(t') \, \mathrm{d}t,' \quad \textit{where} \;\; t' := x - t \\
&= \frac{1}{2\pi} \int_{x-\pi}^{x+\pi} f(x-t) D(t) \, \mathrm{d}t \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_n(t) \, \mathrm{d}t, \quad \textit{since, } t \to f(x-t)D_n(t), \textit{ has period } 2\pi.
\end{aligned}$$

Chapter 8

Conjugate Mappings

Let T be a continuous linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. Then we define $T' : Y^* \to X^*$ by, $T'(y^*) := y^* \circ T$ for each $y^* \in Y^*$, i.e., for each $x \in X$, $[T'(y^*)](x) = y^*(T(x))$. Note that $T'(y^*)$ is indeed a member of X^*.

Similarly, we define $T'' : X^{**} \to Y^{**}$ by, $T''(x^{**}) := x^{**} \circ T'$ for each $x^{**} \in X^{**}$, i.e., for each $y^* \in Y^*$, $[T''(x^{**})](y^*) = [x^{**} \circ T'](y^*) = x^{**}(T'(y^*)) = x^{**}(y^* \circ T)$.

Fact: Let $T : X \to Y$ be a continuous linear operator acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. Then T is an isomorphism if, and only if, T is onto and there exists an $m > 0$ such that $m\|x\| \leqslant \|\!|T(x)|\!\|$ for all $x \in X$.

Proof. First note that T must be one-to-one, since if $x \neq 0$, then $\|\!|T(x)|\!\| \geqslant m\|x\| \neq 0$ i.e., $x \notin \mathrm{Ker}(T)$. Hence, $\mathrm{Ker}(T) = \{0\}$ and so T is one-to-one. Therefore, T^{-1} exists and is linear. We need to show that it is continuous. Consider $y \in Y$. Now, $m\|x\| \leqslant \|\!|T(x)|\!\|$ for all $x \in X$. Therefore, $m\|T^{-1}(y)\| \leqslant \|\!|T(T^{-1}(y))|\!\| = \|\!|y|\!\|$. That is, $\|T^{-1}(y)\| \leqslant M\|\!|y|\!\|$ for all $y \in Y$ where, $M := 1/m$. $\square$

Fact: Let T be a continuous linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. Then T' is one-to-one if, and only if, $\overline{T(X)} = Y$. In particular, if X or Y are finite dimensional, then T' is one-to-one if, and only if, T is onto.

Proof. Suppose that $\overline{T(X)} = Y$ and consider $y^* \in Y^*$ such that $T'(y^*) = 0$, i.e., $y^* \circ T = 0$. Then, for each $x \in X, y^*(T(x)) = 0$, i.e., $y^*|_{T(X)} = 0$. Since y^* is continuous, this implies that $y^* = 0$ on $\overline{T(X)} = Y$. Thus, if $T'(y^*) = 0$, then $y^* = 0$.

Now, suppose T' is one-to-one, but $\overline{T(X)} \neq Y$. Then by Exercise 4.15 there exists a $y^* \in S_{Y^*}$ such that $y^*(T(X)) = \{0\}$. Then $T'(y^*) = 0$; which implies that T^* is not one-to-one. $\square$

Corollary 8.1. *Let T be a continuous linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. Then T'' is one-to-one if, and only if, $\overline{T'(Y^*)} = X^*$.*

Fact: Let T be a continuous linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. Then, $T''|_{\widehat{X}} = \widehat{T}$, where $\widehat{T} : \widehat{X} \to \widehat{Y}$ is defined by, $\widehat{T}(\widehat{x}) := \widehat{T(x)}$.

Proof. $T''(\widehat{x}) = \widehat{x} \circ T'$. Therefore, for any $y^* \in Y^*$,

$$[T''(\widehat{x})](y^*) = [\hat{x} \circ T'](y^*) = \widehat{x}(T'(y^*)) = \widehat{x}(y^* \circ T) = (y^* \circ T)(x) = y^*(T(x)) = \widehat{T(x)}(y^*).$$

Thus, $T''(\widehat{x}) = \widehat{T(x)} = \widehat{T}(\widehat{x})$. $\square$

Corollary 8.2. *Let T be a continuous linear mapping acting between normed linear spaces $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$. If $\overline{T'(Y^*)} = X^*$, then T is one-to-one.*

Warning : The converse is not true! That is, there exist 1-to-1 mappings T such that T' does not have dense range.

Fact: Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ and $(Z, \|\cdot\|_Z)$ be normed linear spaces and suppose $S \in B(X, Y)$ and $T \in B(Y, Z)$. Then $(T \circ S)' \in B(Z^*, X^*)$ and $(T \circ S)' = S' \circ T'$.

Proof. Firstly, $S' \circ T'$ is well defined since $T' \in B(Z^*, Y^*)$ and $S' \in B(Y^*, X^*)$. Now,

$$(T \circ S)'(z^*) = z^* \circ (T \circ S) = (z^* \circ T) \circ S = (T'(z^*)) \circ S = S'(T'(z^*)) = (S' \circ T')(z^*).$$

for any $z^* \in Z^*$. Therefore, $(T \circ S)' = S' \circ T'$. $\square$

Exercise 8.3. *Let $(X, \|\cdot\|)$ be a normed linear space. Show that $(I_X)' = I_{X^*}$, where I_X is the identity mapping on X and I_{X^*} is the identity mapping on X^*.*

Theorem 8.4. *Let $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|))$ be a Banach spaces and let $T : X \to Y$. Then T is an isomorphism if, and only if, $T' : Y^* \to X^*$ is an isomorphism.*

Proof. Suppose T is an isomorphism from X onto Y. Then,

$$(T' \circ (T^{-1})') = (T^{-1} \circ T)' = (I_X)' = I_{X^*}$$

and

$$((T^{-1})' \circ T') = (T \circ T^{-1})' = (I_Y)' = I_{Y^*}.$$

Therefore, $(T')^{-1} = (T^{-1})'$.

Now, suppose that T' is an isomorphism, then in particular, T' is one-to-one. Therefore, $\overline{T(X)} = Y$. Since T'' is an isomorphism there exists an $m > 0$ such that $\|\!|T''(x^{**})|\!\| \geqslant m\|x^{**}\|$. Hence,

$$\|\!|T(x)|\!\| = \|\!|\widehat{T(x)}|\!\| = \|\!|\widehat{T}(\widehat{x})|\!\| = \|\!|T''(\widehat{x})|\!\| \geqslant m\|\widehat{x}\| = m\|x\|.$$

Thus, T is one-to-one, and an isomorphism onto $T(X)$. Since $(X, \|\cdot\|)$ is a Banach space, $T(X)$ is also a Banach space, with the restriction of the norm $\|\!|\cdot|\!\|$ on Y to $T(X)$, and is therefore a closed subspace. Hence, $T(X) = \overline{T(X)} = Y$ and so T is an isomorphism. $\square$

What does T' look like in finite dimensions? Suppose that $(X, \|\cdot\|)$ is a finite dimensional normed linear space.

Let $(e_n)_{n=1}^N$ be a basis for X and for each $1 \leqslant n \leqslant N$, let $e_n^* : X \to \mathbb{K}$ be defined by,

$$e_n^*\left(\sum_{k=1}^N x_k e_k\right) := x_n.$$

In particular, $e_n^*(e_k) = \delta_{nk}$ for each $1 \leqslant k \leqslant N$ and $1 \leqslant n \leqslant N$.

Claim: For each $x^* \in X^*$, $x^* = \sum_{k=1}^N x^*(e_k)e_k^*$.
To see this, observe that

$$x^*(e_n) = \left(\sum_{k=1}^N x^*(e_k)e_k^*\right)(e_n) \quad \text{for all } 1 \leqslant n \leqslant N.$$

Also, if $\sum_{k=1}^N c_k e_k^* = 0$ for some $(c_k)_{k=1}^N \in \mathbb{K}^N$, then for each $1 \leqslant n \leqslant N$,

$$c_n = \left(\sum_{k=1}^N c_k e_k^*\right)(e_n) = 0.$$

Hence $(e_k^*)_{k=1}^N$ is a basis for X^*. In particular, $\dim(X) = \dim(X^*)$.

Now, suppose that both $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot|\!\|)$ are finite dimensional normed linear spaces and $T : X \to Y$ is linear. Let $(e_k)_{k=1}^n$ be a basis for X and $(f_k)_{k=1}^m$ be a basis for Y. Also, let A be the $m \times n$ matrix representation of T with respect to $(e_k)_{k=1}^n$ and $(f_k)_{k=1}^m$ (That is, $[A]_{ij} = i^{\text{th}}$ coordinate of $T(e_j)$ with respect to $(f_k)_{k=1}^m$).

Similarly, let B be the $n \times m$ matrix representation of $T' : Y^* \to X^*$ with respect to $(f_k^*)_{k=1}^m$ and $(e_k^*)_{k=1}^n$ (That is, $[B]_{ij} = i^{\text{th}}$ coordinate of $T'(f_j^*)$ with respect to $(e_k^*)_{k=1}^n$).

What is the relationship between B and A?

Firstly, A is an $m \times n$ matrix and B is an $n \times m$ matrix. Moreover, $[B]_{ij}$ is the i^{th} coordinate of $T'(f_j^*)$ with respect to $(e_k^*)_{k=1}^n$, i.e.,

$$[B]_{ij} = T'(f_j^*)(e_i) = f_j^*(T(e_i));$$

which is the j^{th} coordinate of $T(e_i)$ with respect to $(f_k)_{k=1}^m$, which is $[A]_{ji}$. That is, $[B]_{ij} = [A]_{ji}$. Thus, $B = A^t$. $\square$

Chapter 9

Reflexive Spaces

We shall say that a normed linear space $(X, \|\cdot\|)$ is **reflexive** if $\widehat{X} = X^{**}$.

Fact: If $(X, \|\cdot\|)$ is reflexive, then $(X, \|\cdot\|)$ is a Banach space.

Fact: If $(X, \|\cdot\|)$ is separable and reflexive, then $(X^{**}, \|\cdot\|)$ and $(X^{*}, \|\cdot\|)$ are also separable.

Corollary 9.1. *$(c_0(\mathbb{N}), \|\cdot\|_\infty)$, $(\ell^1(\mathbb{N}), \|\cdot\|_1)$ and $(C[a,b], \|\cdot\|_\infty)$ are not reflexive. Note: we can also deduce that $(\ell^\infty(\mathbb{N}), \|\cdot\|_\infty)$ is not reflexive since closed subspaces of reflexive spaces are reflexive and $(c_0(\mathbb{N}), \|\cdot\|_\infty)$ is a closed subspace of $(\ell^\infty(\mathbb{N}), \|\cdot\|_\infty)$.*

Theorem 9.2 (James' Theorem)**.** *Let $(X, \|\cdot\|)$ be a Banach space. Then $(X, \|\cdot\|)$ is reflexive if, and only if, for each $x^* \in S_{X^*}$ there exists an $x \in S_X$ such that $\|x^*\| = x^*(x)$.*

Theorem 9.3. *All finite dimensional normed linear spaces are reflexive.*

Proof. Let $(X, \|\cdot\|)$ be a finite dimensional normed linear space. Then $\widehat{X}$ (i.e., the natural embedding of X into X^{**}) is a subspace of X^{**}. However,

$$\dim(\widehat{X}) = \dim(X) = \dim(X^*) = \dim(X^{**}).$$

Therefore, $\widehat{X} = X^{**}$ and so X is reflexive. $\square$

In the next exercise we use the following definition. For each $n \in \mathbb{N}$, $e_n^* : \ell^p(\mathbb{N}) \to \mathbb{K}$ is defined by, $e_n^*((x_k)_{k=1}^\infty) := x_n$. It is easy to show that $e_n^* \in \ell^p(\mathbb{N})^*$ and $\|e_n^*\| = 1$ for all $n \in \mathbb{N}$

Exercise 9.4. *Suppose that $1 < p$, $1 < q$ and $1/p + 1/q = 1$. Show that: $(c_n)_{n=1}^\infty \mapsto \sum_{n=1}^\infty c_n e_n^*$ is an isometry from $(\ell^q(\mathbb{N}), \|\cdot\|_q)$ onto $(\ell^p(\mathbb{N})^*, \|\cdot\|)$.*

Theorem 9.5. *$(\ell^p(\mathbb{N}), \|\cdot\|_p)$ is reflexive for each $1 < p < \infty$.*

Proof. As always, $\widehat{\ell^p}(\mathbb{N})$ is a closed subspace of $\ell^p(\mathbb{N})^{**}$. So it is sufficient to show that $\ell^p(\mathbb{N})^{**} \subseteq \widehat{\ell^p}(\mathbb{N})$. To this end, consider $F \in \ell^p(\mathbb{N})^{**}$. Then

$$\begin{aligned} F\left(\sum_{k=1}^{\infty} c_k e_k^*\right) &= F\left(\lim_{n\to\infty} \sum_{k=1}^{n} c_k e_k^*\right) = \lim_{n\to\infty} F\left(\sum_{k=1}^{n} c_k e_k^*\right) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} c_k F(e_k^*) = \sum_{k=1}^{\infty} c_k F\left(e_k^*\right). \end{aligned}$$

Next we show that $(F(e_k^*))_{k=1}^{\infty} \in \ell^p(\mathbb{N})$, i.e., $\sum_{k=1}^{\infty} |F\left(e_k^*\right)|^p < \infty$. For each $n \in \mathbb{N}$, define $x_n^* \in \ell^p(\mathbb{N})^*$ by,

$$x_n^* := \sum_{k=1}^{n} [\operatorname{sign}[F\left(e_k^*\right)] \cdot |F\left(e_k^*\right)|^{p-1}] e_k^*.$$

Then,

$$\sum_{k=1}^{n} |F\left(e_k^*\right)|^p = F(x_n^*) \leqslant \|F\| \cdot \|x_n^*\|.$$

Now,

$$\|x_n^*\| = \left(\sum_{k=1}^{n} \left(|F\left(e_k^*\right)|^{p-1}\right)^q\right)^{1/q} = \left(\sum_{k=1}^{n} |F\left(e_k^*\right)|^p\right)^{1/q}$$

since $(p-1)q = p$. Therefore, for each $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} |F\left(e_k^*\right)|^p \leqslant \|F\| \left(\sum_{k=1}^{n} |F\left(e_k^*\right)|^p\right)^{1/q}.$$

By dividing both sides by $\left(\sum_{k=1}^{n} |F\left(e_k^*\right)|^p\right)^{1/q}$ we get that for each $n \in \mathbb{N}$,

$$\left(\sum_{k=1}^{n} |F\left(e_k^*\right)|^p\right)^{1/p} = \left(\sum_{k=1}^{n} |F\left(e_k^*\right)|^p\right)^{1-(1/q)} \leqslant \|F\| < \infty.$$

Finally, we claim that $\widehat{(F(e_k^*))}_{k=1}^{\infty} = F$. To see this, note that for each $n \in \mathbb{N}$,

$$\widehat{(F(e_k^*))}_{k=1}^{\infty}(e_n^*) = e_n^*\left((F(e_k^*)_{k=1}^{\infty}\right) = F(e_n^*).$$

Since $\overline{\operatorname{span}}\,(e_n^*)_{n=1}^{\infty} = \ell^p(\mathbb{N})^*$ and both $\widehat{(F(e_k^*))}_{k=1}^{\infty}$ and F are continuous linear functionals on $\ell^p(\mathbb{N})^*$, $\widehat{(F\left(e_k^*\right))}_{k=1}^{\infty} = F$. $\square$

Theorem 9.6. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then its dual space is also a Hilbert space and the mapping $x \mapsto x^*$ from H into H^*, defined by, $x^*(y) := \langle y, x\rangle$ for all $y \in H$, is a conjugate linear isometry.*

Proof. From our earlier work on Hilbert spaces we know that the mapping $x \mapsto x^*$ is onto and an isometry. Let us now show that it is conjugate linear. Suppose $x, y \in H$. Then for any $z \in H$,

$$(x+y)^*(z) = \langle z, (x+y)\rangle = \langle z, x\rangle + \langle z, y\rangle = x^*(z) + y^*(z).$$

Therefore $(x+y)^* = x^* + y^*$. Suppose $\lambda \in \mathbb{C}$ and $x \in H$. Then for any $z \in H$,

$$(\lambda x)^*(z) = \langle z, \lambda x\rangle = \overline{\lambda}\langle z, x\rangle = \overline{\lambda} x^*(z).$$

Therefore, $(\lambda x)^* = \overline{\lambda} x^*$. Next, we define an inner product on H^* as follows. For x^* and $y^* \in H^*$ we define

$$\langle x^*, y^*\rangle := \langle y, x\rangle.$$

We need to check that this indeed defines an inner product:

(i): $\langle x^*, x^*\rangle = \langle x, x\rangle = \|x\|^2 \geqslant 0$ and $\langle x^*, x^*\rangle = 0$ if, and only if, $x^* = 0$.

(ii): For any x^*, y^* and z^* in H^*,

$$\begin{aligned}\langle x^* + y^*, z^*\rangle = \langle (x+y)^*, z^*\rangle = \langle z, x+y\rangle &= \langle z, x\rangle + \langle z, y\rangle \\ &= \langle x^*, z^*\rangle + \langle y^*, z^*\rangle.\end{aligned}$$

(iii): For any x^*, z^* in H^* and $\lambda \in \mathbb{C}$,

$$\langle \lambda x^*, z^*\rangle = \langle (\overline{\lambda} x)^*, z^*\rangle = \langle z, \overline{\lambda} x\rangle = \overline{\overline{\lambda}}\langle z, x\rangle = \lambda\langle z, x\rangle = \lambda\langle x^*, z^*\rangle.$$

(iv): For any x^* and z^* in H^*, $\langle x^*, z^*\rangle = \langle z, x\rangle = \overline{\langle x, z\rangle} = \overline{\langle z^*, x^*\rangle}$. Therefore, this defines an inner product on H^*.

We now need to show that the norm generated by this inner product is consistent with the operator norm on H^*. To this end, let $\|x^*\|_H := \sqrt{\langle x^*, x^*\rangle}$ for all $x^* \in H^*$. Therefore,

$$\|x^*\|_H = \sqrt{\langle x^*, x^*\rangle} = \sqrt{\langle x, x\rangle} = \|x\| = \|x^*\|.$$

for all $x^* \in H^*$, since $x \to x^*$ is an isometry. As $(H^*, \|\cdot\|)$ is a dual space, it is also automatically complete. $\square$

Note: it follows from the proof of Theorem 9.6 that the inner product on H^* is given by, $\langle x^*, y^*\rangle = \langle y, x\rangle$.

Corollary 9.7. *Every Hilbert space is reflexive.*

Proof. Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space. It is sufficient to show that $H^{**} \subseteq \widehat{H}$. To this end, consider $F \in H^{**}$. By Theorem 9.6 we know that $F = f^*$ for some $f \in H^*$ and that $f = x^*$ for some $x \in H$. We clam that $\widehat{x} = F$. To see this consider the following. Let $y^* \in H^*$, then

$$F(y^*) = f^*(y^*) = \langle y^*, f\rangle = \langle y^*, x^*\rangle = \langle x, y\rangle = y^*(x) = \widehat{x}(y^*)$$

Since $y^* \in H^*$ was arbitrary, it follows that $F = \widehat{x}$, and so $H^{**} \subseteq \widehat{H}$. $\square$

Adjoint Operators on Hilbert Spaces

Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $\varphi : H \to H^*$ be defined by, $\varphi(x)(z) := \langle z, x\rangle$ for all $z \in H$, (i.e., in terms of the notation from Theorem 9.6, $\varphi(x) = x^*$).

Given a continuous linear operator T on H we can associate with T another continuous linear operator on H, derived from its conjugate T' on H^*, and the mapping $\varphi : H \to H^*$ defined above.

For a continuous linear operator T on a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ we define the **adjoint of** T by, $T^* := \varphi^{-1} \circ T' \circ \varphi$.

Remarks 9.8. *Since φ, φ^{-1} and T' are additive so too is T^*. Since φ and φ^{-1} both are conjugate homogeneous and T' is homogeneous then T^* is homogeneous. Therefore, T^* is linear. As both φ and φ^{-1} as isometries, and in particular continuous, and T' is continuous, it follows that T^* is also continuous. Thus, $T^* \in B(H)$.*

Theorem 9.9. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $T \in B(H)$, then for any $x, z \in H$, $\langle T(z), x\rangle = \langle z, T^*(x)\rangle$. Moreover, if $S \in B(H)$ and $\langle T(x), z\rangle = \langle x, S(z)\rangle$ for all $x, z \in H$, then $S = T^*$.*

Proof. Suppose that $x, z \in H$. Then

$$\begin{aligned}
\langle z, T^*(x)\rangle = \varphi(T^*(x))(z) &= [(\varphi \circ T^*)(x)](z) \\
&= [(\varphi \circ (\varphi^{-1} \circ T' \circ \varphi))(x)](z) \\
&= [((\varphi \circ \varphi^{-1}) \circ T' \circ \varphi)(x)](z) \\
&= [(T' \circ \varphi)(x)](z) \\
&= T'(\varphi(x))(z) \\
&= (\varphi(x))(T(z)) \\
&= \langle T(z), x\rangle.
\end{aligned}$$

Suppose that $S \in B(H)$ and $\langle T(x), z\rangle = \langle x, S(z)\rangle$ for all $x, z \in H$. Fix $z \in H$ and let x be any member of H. Then,

$$\langle x, T^*(z)\rangle = \langle T(x), z\rangle = \langle x, S(z)\rangle.$$

Therefore, $\langle x, T^*(z) - S(z)\rangle = 0$ for all $x \in H$. In particular, if $x := T^*(z) - S(z)$, then $\|T^*(z) - S(z)\|^2 = 0$ and so $T^*(z) = S(z)$. Since $z \in H$ was arbitrary, $S = T^*$. □

Theorem 9.10. *Given a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ the adjoint mapping $T \mapsto T^*$ defined on $B(H)$ has the properties:*

(i) $(S + T)^* = S^* + T^*$ *for any* $S, T \in B(H)$;

(ii) $(\lambda T)^* = \overline{\lambda} T^*$ *for any* $\lambda \in \mathbb{C}$ *and* $T \in B(H)$;

(iii) $(ST)^* = T^* S^*$ *for any* $S, T \in B(H)$;

(iv) $T^{**} = T$ *for any* $T \in B(H)$;

(v) $\|T^*T\| = \|T\|^2$ *for any* $T \in B(H)$.

Proof. The proof of these facts are left as an exercise for the reader. $\square$

Exercise 9.11. *Let H be a Hilbert space. Show that for any $T \in B(H)$, $\|T^*\| = \|T\|$. Also show that $\langle T^*(x), z\rangle = \langle x, T(z)\rangle$ for any $x, z \in H$.*

What does T^* look like in finite dimensions? Suppose that $(H, \langle\cdot,\cdot\rangle)$ is a finite dimensional Hilbert space and $T \in B(H)$. Let $(e_k)_{k=1}^n$ be an orthonormal basis for H and let A be the $n \times n$ matrix representation of T with respect to $(e_k)_{k=1}^n$ (That is, $[A]_{ij} = i^{\text{th}}$ coordinate of $T(e_j)$ with respect to $(e_k)_{k=1}^n$). Similarly, let B be the $n \times n$ matrix representation of T^* with respect to $(e_k)_{k=1}^n$ (That is, $[B]_{ij} = i^{\text{th}}$ coordinate of $T^*(e_j)$ with respect to $(e_k)_{k=1}^n$).

What is the relationship between B and A? Firstly, A and B have the same shape and moreover,

$$[B]_{ij} = \langle T^*(e_j), e_i\rangle = \langle e_j, T(e_i)\rangle = \overline{\langle T(e_i), e_j\rangle} = \overline{[A]}_{ji}.$$

Therefore, $B = (\overline{A})^t$.

In the next example will be working in $L^2[a,b]$. Recall that $(L^2[a,b], \langle\cdot,\cdot\rangle)$ is a Hilbert space, where the inner product $\langle\cdot,\cdot\rangle$ is defined by,

$$\langle f, g\rangle := \int_{[a,b]} f(t)\overline{g(t)}\ \mathrm{d}t \quad \text{for all } f, g \in L^2[a,b].$$

Note also that $\|f\|_2 = \sqrt{\langle f, f\rangle}$ for all $f \in L^2[a,b]$.

Example 9.12. *Let $K \in C_{\mathbb{C}}([a,b] \times [a,b])$. Then the mapping*

$$T : (L^2[a,b], \|\cdot\|_2) \to (L^2[a,b], \|\cdot\|_2)$$

defined by,

$$T(x)(t) := \int_{[a,b]} K(t,s)x(s)\ \mathrm{d}s \qquad \textit{for all } t \in [a,b] \textit{ and all } x \in L^2[a,b]$$

is a member of $B(L^2[a,b])$.

Claim: $S : (L^2[a,b], \langle\cdot,\cdot\rangle) \to (L^2[a,b], \langle\cdot,\cdot\rangle)$ given by,

$$S(x)(s) = \int_{[a,b]} \overline{K(t,s)}x(t)\ \mathrm{d}t \qquad \text{for all } s \in [a,b] \text{ and all } x \in L^2[a,b]$$

is the adjoint of T, i.e., $S = T^*$.

Proof. It is sufficient to show that for every $x, y \in L^2[a,b]$,

$$\langle T(x), y\rangle = \langle x, S(y)\rangle, \quad \text{that is}$$

$$\int_{[a,b]} [T(x)(t)]\overline{y(t)}\ \mathrm{d}t = \int_{[a,b]} x(s)[\overline{S(y)(s)}]\ \mathrm{d}s.$$

Now,

$$\begin{aligned}
\int_{[a,b]} [T(x)(t)]\overline{y(t)}\ \mathrm{d}t &= \int_{[a,b]} \left(\int_{[a,b]} K(t,s)x(s)\ \mathrm{d}s\right)\overline{y(t)}\ \mathrm{d}t \\
&= \int_{[a,b]} \left(\int_{[a,b]} K(t,s)x(s)\overline{y(t)}\ \mathrm{d}s\right)\mathrm{d}t \\
&= \int_{[a,b]\times[a,b]} K(t,s)x(s)\overline{y(t)}\ \mathrm{d}s\mathrm{d}t \\
&= \int_{[a,b]\times[a,b]} K(t,s)x(s)\overline{y(t)}\ \mathrm{d}t\mathrm{d}s \\
&= \int_{[a,b]} x(s)\left(\int_{[a,b]} K(t,s)\overline{y(t)}\ \mathrm{d}t\right)\mathrm{d}s \\
&= \int_{[a,b]} x(s)\overline{\left(\int_{[a,b]} \overline{K(t,s)}y(t)\ \mathrm{d}t\right)}\ \mathrm{d}s \\
&= \int_{[a,b]} x(s)[\overline{S(y)(s)}]\ \mathrm{d}s.
\end{aligned}$$

This complete the proof of the claim. □

Remarks 9.13. *Note that if K is real-valued and symmetric, i.e., $K(s,t) = K(t,s)$ for all $(s,t) \in [a,b]\times[a,b]$, then $T = T^*$. In this case we call T* **self-adjoint**.

Chapter 10

Stone-Weierstrass Theorem

Let (T, τ) be a topological space. We shall denote by $C(T)$ the space of all bounded real-valued continuous functions defined on T. We shall say that a nonempty subset $\mathscr{A}$ of $C(T)$ is an **algebra** if it is a vector subspace of $C(T)$, i.e., closed under pointwise scalar multiplication and pointwise addition, and is also closed under pointwise multiplication, i.e., if $f, g \in \mathscr{A}$, then $f \cdot g \in \mathscr{A}$, where $(f \cdot g)(t) := f(t)g(t)$ for each $t \in T$.

We shall say that a subset L of $C(T)$ is a **lattice** if it is closed under taking pointwise maximums and pointwise minimums, i.e., if $f, g \in L$, then $f \vee g \in L$ and $f \wedge g \in L$, where $(f \vee g)(t) := \max\{f(t), g(t)\}$ for each $t \in T$ and $(f \wedge g)(t) := \min\{f(t), g(t)\}$ for each $t \in T$.

Exercise 10.1. *Let (T, τ) be a topological space and let S be a vector subspace of $C(T)$. Show that S is a lattice if, and only if, $|f| \in S$ for every $f \in S$.*

Exercise 10.2. *Let (T, τ) be a topological space. Show that if $\mathscr{A}$ is a subalgebra of $C(T)$, then the closure of $\mathscr{A}$ in $(C(T), \|\cdot\|_\infty)$ is also a subalgebra of $C(T)$.*

Theorem 10.3. *There exists a sequence of polynomials $(P_n : n \in \mathbb{N})$, without constant terms, defined on $\mathbb{R}$ that converge uniformly on $[-1, 1]$ to the function $g : [-1, 1] \to [0, 1]$ defined by, $g(x) := |x|$ for all $x \in [-1, 1]$.*

Proof. Let us inductively define a sequence $(P_n : n \in \mathbb{N})$ of polynomials by, $P_0(t) := 0$ for all $t \in \mathbb{R}$ and $P_{n+1}(t) := P_n(t) + (1/2)[t^2 - P_n(t)^2]$ for all $t \in \mathbb{R}$. Clearly each P_n is a polynomial and $P_{n+1}(t) = P_n(t) + (1/2)(|t| - P_n(t))(|t| + P_n(t))$ for all $t \in \mathbb{R}$.

We shall prove, by induction, that

$$0 \leqslant |t| - P_n(t) \leqslant 2|t|/(2 + n|t|) \leqslant 2/(2 + n) \text{ for all } -1 \leqslant t \leqslant 1 \text{ and all } n \in \mathbb{N}.$$

Firstly, let us note that the inequality $2|t|/(2+n|t|) \leqslant 2/(2+n)$ for all $-1 \leqslant t \leqslant 1$ follows directly from cross multiplying. Next, let us note that

$$\begin{aligned} |t| - P_{n+1}(t) &= |t| - \big[P_n(t) + (1/2)(|t| - P_n(t))(|t| + P_n(t))\big] \\ &= \big[|t| - P_n(t)\big] - (1/2)(|t| - P_n(t))(|t| + P_n(t)) \\ &= \big[|t| - P_n(t)\big]\big[1 - (1/2)(|t| + P_n(t))\big] \quad \text{for all } n \in \mathbb{N}. \quad (*) \end{aligned}$$

Using equation $(*)$ and the recursive definition of the polynomials P_n we may deduce, via induction, that $0 \leqslant P_n(t) \leqslant |t|$ for all $n \in \mathbb{N}$ and $t \in [-1, 1]$. Indeed, if $0 \leqslant P_n(t) \leqslant |t|$ for all $t \in [-1, 1]$, then $0 \leqslant t^2 - P_n(t)^2$ and so $P_{n+1}(t) = P_n(t) + (1/2)[t^2 - P_n(t)^2] \geqslant 0$.

Note also that if $P_n(t) \leqslant |t|$ and $t \in [-1, 1]$, then $(1/2)[|t| + P_n(t)] \leqslant 1$ and so

$$0 \leqslant \big(1 - (1/2)(|t| + P_n(t))\big).$$

Therefore, if $t \in [-1, 1]$ and $0 \leqslant P_n(t) \leqslant |t|$, then by Equation $(*)$ we have that $0 \leqslant |t| - P_{n+1}(t)$ for all $t \in [-, 1, 1]$. Thus, $P_{n+1}(t) \leqslant |t|$ for all $t \in [-1, 1]$.

Now, since $0 \leqslant P_n(t)$ for all $t \in [-1, 1]$, $1 - (1/2)\big[|t| + P_n(t)\big] \leqslant 1 - (1/2)|t|$. Therefore,

$$\begin{aligned} \big[2 + (n+1)|t|\big]\big[1 - (1/2)(|t| + P_n(t))\big] &\leqslant \big[2 + (n+1)|t|\big]\big[1 - (1/2)|t|\big] \\ &= 2 + (n+1)|t| - (|t|/2)\big[2 + (n+1)|t|\big] \\ &= 2 + n|t| - \big[(n+1)/2\big]|t|^2 \\ &\leqslant 2 + n|t| \quad \text{for all } n \in \mathbb{N} \text{ and } t \in [-1, 1]. \end{aligned}$$

Therefore, by cross multiplying, we get that:

$$\frac{1}{2 + n|t|}\Big[1 - (1/2)(|t| + P_n(t))\Big] \leqslant \frac{1}{2 + (n+1)|t|} \quad \text{for all } n \in \mathbb{N} \text{ and } t \in [-1, 1]$$

Then, by multiplying through by $2|t|$, we get that:

$$\frac{2|t|}{2 + n|t|}\Big[1 - (1/2)(|t| + P_n(t))\Big] \leqslant \frac{2|t|}{2 + (n+1)|t|} \quad (**)$$

for all $n \in \mathbb{N}$ and $t \in [-1, 1]$. The inequality $|t| - P_n(t) \leqslant 2|t|/(2 + n|t|)$ now follows from induction by applying the inequality $(**)$ to equation $(*)$. $\square$

Theorem 10.4. *Let (T, τ) be a topological space and let $\mathscr{A}$ be a subalgebra of $C(T)$. Then the closure of $\mathscr{A}$ in $(C(T), \|\cdot\|_\infty)$, denoted $\overline{\mathscr{A}}$, is a sublattice of $C(T)$.*

Proof. By Exercise 10.2, $\overline{\mathscr{A}}$ is a subalgebra of $C(T)$, and in particular, a subspace of $C(T)$. So by Exercise 10.1 we need only show that $|f| \in \overline{\mathscr{A}}$. In fact, because $\overline{\mathscr{A}}$ is homogeneous, we need only show that $|f| \in \overline{\mathscr{A}}$, whenever $f \in \overline{\mathscr{A}}$ and $\|f\|_\infty = 1$.

Now, from Theorem 10.3 there exist polynomials $(P_n : n \in \mathbb{N})$, without constant terms, on $\mathbb{R}$ such that

$$|f| = \lim_{n\to\infty}(P_n \circ f)$$

in $(C(T), \|\cdot\|_\infty)$. Therefore, since $(P_n \circ f) \in \overline{\mathscr{A}}$ for all $n \in \mathbb{N}, |f| \in \overline{\mathscr{A}}$. $\square$

Let (T, τ) be a topological space and let S be a subset of $C(T)$. We shall say that S has the **2-point approximation property** if for every $f \in C(T), x, y \in T$ and $\varepsilon > 0$ there exists an $s \in S$ such that $|s(x) - f(x)| < \varepsilon$ and $|s(y) - f(y)| < \varepsilon$.

Theorem 10.5 (Stone-Weierstrass Theorem)**.** *Let (T, τ) be a compact space and let L be a sublattice of $C(T)$. If L possesses the 2-point approximation property, then $\overline{L} = C(T)$.*

Proof. Let $f \in C(T)$ and $\varepsilon > 0$. It will be sufficient to show that there exists a $g \in L$ such that $\|f - g\| < \varepsilon$. Fix $x \in T$. For each $y \in T$ there exists an open neighbourhood U_y^x of y and an element $g_y^x \in L$ such that $g_y^x(x) < f(x) + \varepsilon$ and $f(t) - \varepsilon < g_y^x(t)$ for all $t \in U_y^x$. Let $\{U_{y_j}^x : 1 \leqslant j \leqslant n\}$ be a finite subcover of $\{U_y^x : y \in T\}$ and let $g_x : T \to \mathbb{R}$ be defined by,

$$g_x(t) := \max_{1 \leqslant j \leqslant n} g_{y_j}^x(t)$$

i.e., $g_x = \bigvee_{1 \leqslant j \leqslant n} g_{y_j}^x \in L$. Then $g_x(x) < f(x) + \varepsilon$ while $f(t) - \varepsilon < g_x(t)$ for all $t \in T$.

We now consider the family of functions $\{g_x : x \in T\}$. For each $x \in T$ there exists an open neighbourhood V_x of x such that $g_x(t) < f(t) + \varepsilon$ for all $t \in V_x$. Let $\{V_{x_j} : 1 \leqslant j \leqslant m\}$ be a finite subcover of $\{V_x : x \in T\}$ and define $g : T \to \mathbb{R}$ by,

$$g(t) := \min_{1 \leqslant j \leqslant m} g_{x_j}(t)$$

i.e., $g = \bigwedge_{1 \leqslant j \leqslant m} g_{x_j} \in L$. It is easily seen that $|g(t) - f(t)| < \varepsilon$ for each $t \in T$ and so $\|g - f\|_\infty < \varepsilon$. $\square$

Corollary 10.6. *Let (T, τ) be a compact space and let $\mathscr{A}$ be a subalgebra of $C(T)$. If $\mathscr{A}$ possesses the 2-point approximation property, then $C(T) = \overline{\mathscr{A}}$.*

Proof. By Theorem 10.4, $\overline{\mathscr{A}}$ is a lattice. Since $\mathscr{A} \subseteq \overline{\mathscr{A}}$, $\overline{\mathscr{A}}$ clearly possesses the 2-point approximation property. Therefore, by Theorem 10.5, $C(T) = \overline{\overline{\mathscr{A}}} = \overline{\mathscr{A}}$. $\square$

Corollary 10.7. *Let (T, τ) be a compact space and let $\mathscr{A}$ be a subalgebra of $C(T)$ that contains all the constant functions and separates the points of T (i.e., if $x \neq y \in T$, then there exists an $f \in \mathscr{A}$ such that $f(x) \neq f(y)$), then $C(T) = \overline{\mathscr{A}}$.*

Proof. If $\mathscr{A}$ contains all the constant functions and separates the point of T, then $\mathscr{A}$ has the 2-point approximation property. The result then follows from Corollary 10.6. $\square$

Let (T, τ) be a topological space. We shall denote by, $C_{\mathbb{C}}(T)$ the space of all bounded complex-valued continuous functions defined on T. We shall say that a subalgebra $\mathscr{A}$ of $C_{\mathbb{C}}(T)$ is **self-adjoint** if $\overline{f} \in \mathscr{A}$ whenever $f \in \mathscr{A}$, where $\overline{f} : T \to \mathbb{C}$ is defined by, $\overline{f}(t) := \overline{f(t)}$ for each $t \in T$.

Theorem 10.8. *Let (T, τ) be a compact space and let $\mathscr{A}$ be a self-adjoint subalgebra of $C_{\mathbb{C}}(T)$ that contains all the constant functions and separates the points of T, then $C_{\mathbb{C}}(T) = \overline{\mathscr{A}}$.*

Proof. The proof of this is left as an exercise for the reader. $\square$

Theorem 10.9. *Let (X,τ) and (Y,τ') be compact spaces. Then for each $h \in C(X\times Y)$ and $\varepsilon > 0$ there exist $(f_j)_{j=1}^n$ in $C(X)$ and $(g_j)_{j=1}^n$ in $C(Y)$ such that*

$$\left| h(x,y) - \sum_{j=1}^{n} f_j(x)g_j(y) \right| < \varepsilon \quad \textit{for all } (x,y) \in X\times Y.$$

Proof. The proof of this is left as an exercise for the reader. □

Theorem 10.10. *The set $\left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$ is an orthonormal basis for the Hilbert space $(L^2[0,2\pi], \langle\cdot,\cdot\rangle)$.*

Proof. We give here only an outline.

(i) First note that $\left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$ is an orthonormal basis if, and only if, $L^2[0,2\pi] = \overline{\operatorname{span}} \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$;

(ii) Justify the fact that $L^2[0,2\pi] = \overline{\operatorname{span}} \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$ if, and only if, $C^*_{\mathbb{C}}[0,2\pi] \subseteq \overline{\operatorname{span}} \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$, where $C^*_{\mathbb{C}}[0,2\pi] := \{ f \in C_{\mathbb{C}}[0,2\pi] : f(0) = f(2\pi) \}$;

(iii) Let $\mathscr{A}$ be the algebra generated by the set $\left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$. Show that $\mathscr{A} = \operatorname{span} \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$;

(iv) Show that $\mathscr{A}$ is a self-adjoint algebra;

(v) Adapt the proof of the Stone-Weierstrass Theorem to show that

$$C^*_{\mathbb{C}}[0,2\pi] = \overline{\operatorname{span}} \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$$

considered in $(C^*_{\mathbb{C}}[0,2\pi], \|\cdot\|_\infty)$;

(vi) Hence deduce that $C^*_{\mathbb{C}}[0,2\pi] \subseteq \overline{\operatorname{span}} \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} : k \in \mathbb{Z} \right\}$ when considered in $(L^2[0,2\pi], \|\cdot\|_2)$.

This completes the proof. □

Chapter 11

Arzelà-Ascoli Theorem

A subset T of a metric space (X, d) is called **totally bounded** if for each $\varepsilon > 0$ there exists a finite subset F_ε of X such that $T \subseteq \bigcup\{B[x; \varepsilon] : x \in F_\varepsilon\}$.

Theorem 11.1. *Let (X, d) be a complete metric space and let K be a closed and totally bounded subset of (X, d). Then K is compact.*

Proof. Let $(x_n : n \in \mathbb{N})$ be a sequence in K. We need to show that $(x_n : n \in \mathbb{N})$ possesses a subsequence that is Cauchy. For each $n \in \mathbb{N}$, let $\{C_j^n : 1 \leqslant j \leqslant N_n\}$ be a finite cover of K by sets with diameter less than $1/n$. Note: this is possible since K is totally bounded. We shall inductively construct infinite subsets $\{J_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ such that:

(i) $J_{n+1} \subseteq J_n$ for all $n \in \mathbb{N}$;

(ii) for each $n \in \mathbb{N}$ there exists a $j_n \in \{1, 2, \ldots, N_n\}$ such that $x_k \in C_{j_n}^n$ for all $k \in J_n$.

The construction of these sets is left as an exercise for the reader. Next, we may define $(n_k : k \in \mathbb{N})$ such that:

(i) $n_k < n_{k+1}$ for all $k \in \mathbb{N}$;

(ii) $n_k \in J_k$ for all $k \in \mathbb{N}$.

Now, since $J_{n+1} \subseteq J_n$ for all $n \in \mathbb{N}$ and $n_k \in J_k$ for all $k \in \mathbb{N}$ we have that for each $N \in \mathbb{N}$, $n_k \in J_N$ for all $k \geqslant N$. Therefore, for each $N \in \mathbb{N}$, $\operatorname{diam}\{x_{n_k} : k \geqslant N\} < 1/N$. Hence $(x_{n_k} : k \in \mathbb{N})$ is a Cauchy sequence. $\square$

Corollary 11.2. *Let $(X, \|\cdot\|)$ be a Banach space and let K be a closed subset of $(X, \|\cdot\|)$. Then K is compact if for each $\varepsilon > 0$ there exists a compact subset C_ε of $(X, \|\cdot\|)$ such that $K \subseteq C_\varepsilon + \varepsilon B_X$.*

Let (T, τ) be a topological space. We shall say that a subset F of $C(T)$ is **equicontinuous on** T if for every $\varepsilon > 0$ and every $t \in T$ there exists a neighbourhood $U(t, \varepsilon)$ of t such that $|f(t') - f(t)| < \varepsilon$ for all $t' \in U(t, \varepsilon)$ and all $f \in F$.

Theorem 11.3 (Arzelà-Ascoli Theorem)**.** *Let (T,τ) be a compact space and let K be a nonempty subset of $C(T)$. Then $\overline{K}$ is compact in $(C(T),\|\cdot\|_\infty)$ if, and only if, K is bounded and equicontinuous on T.*

Proof. Suppose that $\overline{K}$ is compact. Consider the function $d:\overline{K}\to[0,\infty)$ defined by, $d(f):=\|f\|_\infty$. Then d is continuous on $\overline{K}$ (since $|d(f)-d(g)|\leqslant\|f-g\|_\infty$) and hence bounded above by some $M>0$. Then $\|f\|_\infty=d(f)\leqslant M$ for all $f\in\overline{K}$ (i.e., K is bounded).

We will now show that K is equicontinuous on T. To see this, consider $t\in T$ and $\varepsilon>0$. Since $\overline{K}$ is compact there exists a finite set $(f_n)_{n=1}^N$ in $\overline{K}$ such that $K\subseteq\bigcup_{n=1}^N B(f_n,\varepsilon/3)$. For each $1\leqslant n\leqslant N$, choose a neighbourhood $U(t,n,\varepsilon)$ of t such that $|f_n(t')-f_n(t)|<\varepsilon/3$ for all $t'\in U(t,n,\varepsilon)$ and let $U(t,\varepsilon):=\bigcap_{n=1}^N U(t,n,\varepsilon)$. Let $f\in K$ and let $t'\in U(t,\varepsilon)$. Then choose $k\in\{1,2,\ldots,N\}$ so that $\|f-f_k\|_\infty<\varepsilon/3$. Thus,

$$\begin{aligned}|f(t)-f(t')| &\leqslant |f(t)-f_k(t)|+|f_k(t)-f_k(t')|+|f_k(t')-f(t')|\\ &\leqslant 2\|f-f_k\|_\infty+|f_k(t)-f_k(t')|\\ &< 2\cdot\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

Hence, K is equicontinuous.

Converse direction. Since we know that $(C(T),\|\cdot\|_\infty)$ is complete and $\overline{K}$ is closed it is sufficient to show that $\overline{K}$ is totally bounded. Thus, let us fix $\varepsilon>0$. For each $x\in T$ there exists an open neighbourhood V_x of x such that $|f(y)-f(x)|<\varepsilon$ for all $y\in V_x$ and all $f\in K$. Since T is compact and $\{V_x:x\in T\}$ is an open cover of T there exists a finite subcover $\{V_{x_1},V_{x_2},\ldots,V_{x_n}\}$ of T. Now, $\{f(x_i):i\in\{1,2,\ldots,n\},f\in K\}$ is bounded in $\mathbb{R}$. Therefore there exist real numbers $\{y_1,y_2,\ldots,y_m\}$ such that

$$\{f(x_i):i\in\{1,2,\ldots,n\},f\in K\}\subseteq B(y_1,\varepsilon)\cup B(y_2,\varepsilon)\cup\cdots\cup B(y_m,\varepsilon).$$

Let $\pi:\{1,2,\ldots,n\}\to\{1,2,\ldots m\}$ be a function. Then define,

$$S_\pi:=\{f\in K:f(x_i)\in B(y_{\pi(i)},\varepsilon)\text{ for all }1\leqslant i\leqslant n\}.$$

Note that

$$\left\{S_\pi:\pi\in\{1,2,\ldots,m\}^{\{1,2,\ldots,n\}}\right\}$$

is a cover of K. Next, we will show that each S_π has diameter at most 4ε. To this end, let $\pi\in\{1,2,\ldots,m\}^{\{1,2,\ldots,n\}}$, let $f,f',\in S_\pi$ and let $x\in T$. Then there exists an $i\in\{1,2,\ldots,n\}$ such that $x\in V_{x_i}$. Thus,

$$|f(x)-f'(x)|\leqslant|f(x)-f(x_i)|+|f(x_i)-y_{\pi(i)}|+|y_{\pi(i)}-f'(x_i)|+|f'(x_i)-f'(x)|<4\varepsilon.$$

Since $x\in T$ was arbitrary it follows that $\|f-f'\|_\infty\leqslant4\varepsilon$, and since $f,f'\in S_\pi$ were also arbitrary, we have that $\|\cdot\|_\infty-\operatorname{diam}(S_\pi)\leqslant4\varepsilon$. Hence, K can be covered with at most m^n closed balls of radius 4ε. Thus, $\overline{K}$ can also be covered with at most m^n closed balls of radius 4ε, as a finite union of closed sets is again closed. This completes the proof. □

Exercise 11.4. *Prove the following complex-valued version of the Arzelá-Ascoli Theorem: Let* (T, τ) *be a compact space and let* K *be a nonempty subset of* $C_{\mathbb{C}}(T)$*. Then* $\overline{K}$ *is compact in* $(C_{\mathbb{C}}(T), \|\cdot\|_\infty)$ *if, and only if,* K *is bounded and equicontinuous on* T*.*

Exercise 11.5. *Let* K *be a subset of a complete metric space* (X, d)*. Show that* $\overline{K}$ *is compact if, and only if, every sequence in* K *has a Cauchy subsequence.*

Exercise 11.6. *Let* $(X, \|\cdot\|)$ *and* $(Y, \|\!|\cdot|\!\|)$ *be normed linear spaces and suppose that* $T \in B(X, Y)$*. Show that if* $\overline{T(B_X)}$ *is a compact subset of* $(Y, \|\!|\cdot|\!\|)$*, then* $\overline{T'(B_{Y^*})}$ *is a compact subset of* $(X^*, \|\cdot\|)$*.*

Hint: In light of Exercise 11.5, to prove Exercise 11.6 we need only show that every sequence in $T'(B_{Y^*})$ possesses a Cauchy subsequence. On the other hand, if we consider $K := \{y^*|_{\overline{T(B_X)}} : y^* \in B_{Y^*}\}$ as a subset of $(C(\overline{T(B_X)}), \|\cdot\|_\infty)$, then one should be able to show that $\overline{K}$ is compact, by appealing to the Arzelà-Ascoli Theorem.

The result in Exercise 11.6 is called "Schauder's Theorem".

Chapter 12

Banach Algebras

An **algebra** over a field $\mathbb{K}$ is a vector space A over $\mathbb{K}$ with a **multiplication** operation $(a, b) \in A \times A \mapsto ab \in A$ such that:

(i) $x(yz) = (xy)z$ for all $x, y, z \in A$;

(ii) $x(y + z) = xy + xz$ and $(y + z)x = yx + zx$ for all $x, y, z \in A$;

(iii) $\alpha(xy) = (\alpha x)y = x(\alpha y)$ for scalars $\alpha \in \mathbb{K}$ and $x, y \in A$.

In this course all algebras will be over the field of complex numbers. An algebra need not have a multiplicative identity element, i.e., an element $e \in A \setminus \{0\}$ such that $ea = ae = a$ for all $a \in A$. If it does have one, then it can be shown to be unique and we will denote it by $\mathbf{1}_A$. We call $\mathbf{1}_A$ the **identity of** A and we say that A is an **algebra with identity** if A is an algebra that possesses an identity element.

Example 12.1. *Let $M_n(\mathbb{C})$ denote the set of all $n \times n$ matrices over $\mathbb{C}$. Then $M_n(\mathbb{C})$ with the operations of matrix addition and matrix multiplication is an algebra with identity.*

A **Banach algebra** is a Banach space $(A, \|\cdot\|)$ over $\mathbb{C}$ which is also an algebra over $\mathbb{C}$ and in which the norm is related to multiplication by the following inequality $\|ab\| \leqslant \|a\|\|b\|$ for all $a, b \in A$. In this case we say that the norm is **submultiplicative**.

A Banach algebra $(A, \|\cdot\|)$ need not have a multiplicative identity, but if it does and it satisfies $\|\mathbf{1}_A\| = 1$, then we call it a **unital Banach algebra** or else a **Banach algebra with identity**.

Example 12.2. *Some examples of unital Banach algebras*

(i) The space $(C_{\mathbb{C}}(K), \|\cdot\|_\infty)$ of all complex-valued continuous functions defined on a compact space K, with scalar multiplication, addition and multiplication defined pointwise is a unital Banach algebra. The multiplicative identity is the function that maps every element of K to 1.

(ii) Let $D := \{z \in \mathbb{C} : |z| \leqslant 1\}$ and let $A(D)$ be the subset of $(C_{\mathbb{C}}(D), \|\cdot\|_\infty)$ consisting of all the functions that are analytic on $\{z \in \mathbb{C} : |z| < 1\}$. This is called the **disc algebra***. Again the multiplicative identity is the function that maps every element of D to 1.*

(iii) If $(X, \|\cdot\|)$ is a nontrivial Banach space over $\mathbb{C}$, then $(B(X), \|\cdot\|)$ is a unital Banach algebra, with scalar multiplication and addition defined pointwise and multiplication defined by composition, i.e., if $S, T \in B(X)$, then $ST := S \circ T$. The multiplicative identity in this case is the identity mapping on X.

(iv) Let $(G, \cdot)$ be a group with identity e and let

$$\ell^1(G) := \{f \in \mathbb{C}^G : \sum_{g \in G} |f(g)| < \infty\},$$

with scalar multiplication and addition defined pointwise. For $f, g \in \ell^1(G)$ we define the **convolution** *of f and g to be the function $f * g : G \to \mathbb{C}$ defined by,*

$$(f * g)(x) := \sum_{y \in G} f(xy^{-1})g(y).$$

Then $(\ell^1(G), \|\cdot\|_1)$ is a unital Banach algebra. The identity element is the function $\mathbf{1} : G \to \{0, 1\}$ defined by, $\mathbf{1}(x) := 1$ if, and only if, $x = e$.

Proof. (i) We already know that $(C_{\mathbb{C}}(K), \|\cdot\|_\infty)$ is a Banach space and that $C_{\mathbb{C}}(K)$ is closed under pointwise multiplication. Further, if $f, g \in C_{\mathbb{C}}(K)$ and $k \in K$, then $|(fg)(k)| = |f(k)||g(k)| \leqslant \|f\|_\infty \|g\|_\infty$ and so $\|fg\|_\infty \leqslant \|f\|_\infty \|g\|_\infty$. Note also that $\|\mathbf{1}\|_\infty = 1$, where $\mathbf{1} : K \to \mathbb{C}$ is defined by, $\mathbf{1}(k) = 1$ for all $k \in K$.

(ii) It is easy to verify that $A(D)$ is a subalgebra of $C_{\mathbb{C}}(D)$ with identity element $\mathbf{1}$. It also follows, for free, since $A(D)$ is a subset of $C_{\mathbb{C}}(D)$ that the norm is submultiplicative and $\|\mathbf{1}\|_\infty = 1$. It remains to show that $A(D)$ is a closed subalgebra of $C_{\mathbb{C}}(D)$. Suppose that $(f_n : n \in \mathbb{N})$ is a sequence in $A(D)$ converging to f in $(C_{\mathbb{C}}(D), \|\cdot\|_\infty)$. Now suppose that Γ is a simple closed contour with length L lying in D, then

$$\left| \int_\Gamma f_n(z)\ \mathrm{d}z - \int_\Gamma f(z)\ \mathrm{d}z \right| = \left| \int_\Gamma (f_n - f)(z)\ \mathrm{d}z \right| \leqslant \|f_n - f\|_\infty L,$$

and thus $\int_\Gamma f_n(z)\ \mathrm{d}z \to \int_\Gamma f(z)\ \mathrm{d}z$. By Cauchy's Theorem we have that $\int_\Gamma f_n(z)\ \mathrm{d}z = 0$ for all $n \in \mathbb{N}$, hence $\int_\Gamma f(z)\ \mathrm{d}z = 0$. Morera's Theorem then implies that f is analytic on $\{z \in \mathbb{C} : |z| < 1\}$. Thus, $f \in A(D)$.

(iii) The only interesting feature here to check is that for any $S, T \in B(X)$, $\|ST\| \leqslant \|S\|\|T\|$. To see this, let $x \in X$. Then

$$\|(ST)(x)\| = \|S(T(x))\| \leqslant \|S\|\|T(x)\| \leqslant \|S\|\|T\|\|x\|.$$

Since $x \in X$ was arbitrary it follows that $\|ST\| \leqslant \|S\|\|T\|$.

(iv) This is an important example, called the **group algebra** of G, so we will take the opportunity to verify a couple of the axioms to show that $\ell_1(G)$, endowed with the convolution, really is a unital Banach algebra. Specifically, we will show that $\|f * g\|_1 \leqslant \|f\|_1\|g\|_1$ for all $f, g \in \ell_1(G)$ and $\|\mathbf{1}\|_1 = 1$. Of course we do already know that $(\ell_1(G), \|\cdot\|_1)$ is a Banach space.

Let $f, g \in \ell_1(G)$, then

$$\begin{aligned}
\|f * g\|_1 &= \sum_{x\in G} |(f*g)(x)| \\
&= \sum_{x\in G} \left| \sum_{y\in G} f(xy^{-1})g(y) \right| \\
&\leqslant \sum_{x\in G}\sum_{y\in G} |f(xy^{-1})||g(y)| \quad \text{by the triangle inequality} \\
&= \sum_{y\in G}\sum_{x\in G} |f(xy^{-1})||g(y)| \quad \text{swap the order of summation} \\
&= \sum_{y\in G} |g(y)| \left(\sum_{x\in G} |f(xy^{-1})| \right) \\
&= \sum_{y\in G} |g(y)| \left(\sum_{z\in G} |f(z)| \right) \quad \text{since } G = Gy^{-1} \\
&= \sum_{y\in G} |g(y)| \|f\|_1 = \|f\|_1 \sum_{y\in G} |g(y)| = \|f\|_1 \|g\|_1.
\end{aligned}$$

Note also that $\|\mathbf{1}\|_1 = \sum_{x\in G} |\mathbf{1}(x)| = \mathbf{1}(e) = 1.$ $\square$

Exercise 12.3. *Let $(G, \cdot)$ be a group. Show that the convolution operation on $\ell_1(G)$ is associative. Hint: Show that for all $f, g, h \in \ell_1(G)$ and all $x \in G$*

$$((f*g)*h)(x) = \sum \{ f(a)g(b)h(c) : (a,b,c) \in G^3 \text{ and } x = abc \} = (f*(g*h))(x).$$

Note also that for every $x \in G$, $\sum\{|f(a)g(b)h(c)| : (a,b,c) \in G^3 \text{ and } x = abc\} < \infty$.

Finally, note that $\pi : G \to \ell_1(G)$, defined by, $[\pi(g)](x) = 1$ if $x = g$ and $[\pi(g)](x) = 0$ if $x \neq g$, is a group monomorphism from $(G, \cdot)$ into $(\ell_1(G), *)$.

Theorem 12.4. *Every unital Banach algebra is isometrically isomorphic to a unital subalgebra of $B(X)$, for some Banach space $(X, \|\cdot\|)$.*

Proof. Let $(A, \|\cdot\|)$ be a unital Banach algebra. Consider the mapping $M : A \to B(A)$ defined by, $M(a)(x) := ax$ for all $x \in A$. One can verify that M is indeed an isometric isomorphism and that $M(A)$ is a unital Banach subalgebra of $B(A)$. $\square$

An element a of a unital algebra A is **invertible** if there exists an element $b \in A$ such that $ab = ba = \mathbf{1}_A$. Note that if $ab = ba = \mathbf{1}_A$ and $ac = ca = \mathbf{1}_A$, then $b = c$. Simply note that $b = b\mathbf{1}_A = b(ac) = (ba)c = \mathbf{1}_A c = c$. Any element $b \in A$ such that $ab = ba = \mathbf{1}_A$ is called an **inverse** of a and by our previous argument we see that the inverse of a is unique. Hence, if $a \in A$ is invertible, then we can denote its inverse by a^{-1}.

Basic facts: Let $(A, \|\cdot\|)$ be a unital Banach algebra, then

(i) If $A^{-1} := \{a \in A : a^{-1} \text{ exists}\}$, then $(A^{-1}, \cdot)$ is a group, called the **group of units** or **group of regular elements**.

(ii) $(x, y) \mapsto x \cdot y$ is jointly continuous, that is, if $\lim\limits_{n\to\infty} x_n = x$ and $\lim\limits_{n\to\infty} y_n = y$, then $\lim\limits_{n\to\infty}(x_n \cdot y_n) = x \cdot y$.

(iii) If $x, y \in A^{-1}$, then $(xy)^{-1} = y^{-1}x^{-1}$ and if $\lambda \neq 0$, then $\lambda x \in A^{-1}$ and
$$(\lambda x)^{-1} = \lambda^{-1}x^{-1}.$$

(iv) If $xy = yx$, then $xy \in A^{-1}$ if, and only if, both $x \in A^{-1}$ and $y \in A^{-1}$.

(v) If $a \in A^{-1}$, then the mapping $T_a : A \to A$ defined by, $T_a(x) := ax$ for all $x \in A$ is a homeomorphism, i.e., T_a is one-to-one and onto and both T_a and T_a^{-1} are continuous.

(vi) If $x, y \in A^{-1}$, then $y^{-1} - x^{-1} = x^{-1}(x - y)y^{-1} = y^{-1}(x - y)x^{-1}$.

Exercise 12.5. *This exercise concerns inverses.*

(i) Let K be a nonempty compact space. Show that an element f of $C_{\mathbb{C}}(K)$ is invertible if, and only if, 0 is not in the image of f, i.e., if $0 \notin f(K)$.

(ii) Show that an element f of $A(D)$ is invertible if, and only if, 0 is not in the image of f, i.e., if $0 \notin f(D)$.

(iii) Let $(X, \|\cdot\|)$ be a Banach space. Show that $S \in B(X)$ is invertible if, and only if, S is a bijection.

(iv) Let $A \in M_n(\mathbb{C})$. Show that A is invertible if, and only if, Ker$(A) = \{0\}$.

Theorem 12.6. *Let $(A, \|\cdot\|)$ be a Banach algebra. Then for each $x \in A$,*
$$\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} \quad \textit{exists}$$
and
$$\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = \inf\{\|x^n\|^{\frac{1}{n}} : n \in \mathbb{N}\}.$$

Proof. Clearly the result is true if $x = 0$, so we shall consider the case when $0 < \|x\|$. First note that $\|x^n\| \leqslant \|x\|^n$ for all $n \in \mathbb{N}$ and so $\|x^n\|^{\frac{1}{n}} \leqslant \|x\|$ for all $n \in \mathbb{N}$. Therefore, $\limsup\limits_{n\to\infty} \|x^n\|^{\frac{1}{n}}$ exists. Hence it will be sufficient to show that if
$$M := \inf\{\|x^n\|^{\frac{1}{n}} : n \in \mathbb{N}\}, \text{ then } M = \limsup_{n\to\infty} \|x^n\|^{\frac{1}{n}}.$$

To this end, let $\varepsilon > 0$ and choose $m \in \mathbb{N}$ such that $\|x^m\|^{\frac{1}{m}} < M + \varepsilon$. Then for each $n \in \mathbb{N}$, there exists $q_n \in \mathbb{N}$ and $0 \leqslant r_n < m$ such that $n = q_n m + r_n$. Thus,
$$\begin{aligned}
M \leqslant \|x^n\|^{\frac{1}{n}} &= \|x^{q_n m + r_n}\|^{\frac{1}{n}} \\
&\leqslant \|x^{q_n m}\|^{\frac{1}{n}} \cdot \|x^{r_n}\|^{\frac{1}{n}} \\
&\leqslant \|x^m\|^{\frac{q_n}{n}} \cdot \|x\|^{\frac{r_n}{n}} \\
&= \left(\|x^m\|^{\frac{1}{m}}\right)^{\frac{q_n m}{n}} \cdot \|x\|^{\frac{r_n}{n}} \\
&\leqslant (M + \varepsilon)^{\frac{q_n m}{n}} \cdot \|x\|^{\frac{r_n}{n}} \quad \text{for all } n \in \mathbb{N}.
\end{aligned}$$

Therefore, since $\lim\limits_{n\to\infty} \dfrac{r_n}{n} = 0$ and $\lim\limits_{n\to\infty} \dfrac{q_n m}{n} = \lim\limits_{n\to\infty} 1 - \dfrac{r_n}{n} = 1$,

$$M \leqslant \limsup_{n\to\infty} \|x^n\|^{\frac{1}{n}} \leqslant \limsup_{n\to\infty} (M+\varepsilon)^{\frac{q_n m}{n}} \|x\|^{\frac{r_n}{n}} = \lim_{n\to\infty} (M+\varepsilon)^{\frac{q_n m}{n}} \|x\|^{\frac{r_n}{n}} = M + \varepsilon.$$

Thus, since ε was arbitrary, $\limsup\limits_{n\to\infty} \|x^n\|^{\frac{1}{n}} = M$. □

Exercise 12.7. *Let $(A, \|\cdot\|)$ be a Banach algebra. Show that if $x \in A$ and*

$$\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = \|x\|,$$

then $\|x^n\| = \|x\|^n$ for all $n \in \mathbb{N}$.

Theorem 12.8. *Let $(A, \|\cdot\|)$ be a unital Banach algebra. If $x \in A$ and $\lim\limits_{n\to\infty} \|x^n\|^{\frac{1}{n}} < 1$, then $(\mathbf{1}_A - x) \in A^{-1}$ and $(\mathbf{1}_A - x)^{-1} = \mathbf{1}_A + \sum_{n\in\mathbb{N}} x^n$.*

Proof. For each $n \in \mathbb{N}$, let

$$s_n := \mathbf{1}_A + \sum_{k=1}^{n} x^k.$$

Then notice that by the "Root Test" for convergence, $\sum\limits_{k=0}^{\infty} \|x^k\| < \infty$. Therefore, since $(A, \|\cdot\|)$ is a Banach space

$$\mathbf{1}_A + \sum_{k\in\mathbb{N}} x^k = \lim_{n\to\infty} s_n \quad \text{exists.}$$

Moreover,

$$(\mathbf{1}_A - x)s_n = \sum_{k=0}^{n} x^k - \sum_{k=1}^{n+1} x^k = (\mathbf{1}_A - x^{n+1}) = \sum_{k=0}^{n} x^k - \sum_{k=1}^{n+1} x^k = s_n(\mathbf{1}_A - x).$$

Therefore,

$$\begin{aligned}(\mathbf{1}_A - x)(\mathbf{1} + \sum_{k=1}^{\infty} x^k) = \lim_{n\to\infty} (\mathbf{1}_A - x)s_n &= \lim_{n\to\infty} (\mathbf{1}_A - x^{n+1}) \\ &= \mathbf{1}_A \\ &= \lim_{n\to\infty} s_n(\mathbf{1}_A - x) = (\mathbf{1}_A + \sum_{k=1}^{\infty} x^k)(\mathbf{1}_A - x).\end{aligned}$$

Thus, $(\mathbf{1}_A - x)^{-1} = \mathbf{1}_A + \sum_{k\in\mathbb{N}} x^k$. □

Remarks 12.9. *Given a unital Banach algebra* $(A, \|\cdot\|)$ *and an element* $x \in A$ *the previous theorem shows that* $(\mathbf{1}_A - \lambda x)$ *is regular provided* $\lim_{n\to\infty} \|(\lambda x)^n\|^{\frac{1}{n}} < 1$*; that is, provided that:*

$$0 \leqslant |\lambda| < 1/(\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}), \textit{ if } \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} \neq 0 \textit{ and for all } \lambda \in \mathbb{C} \textit{ if, } \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = 0.$$

For any such λ *in this range,*

$$(\mathbf{1}_A - \lambda x)^{-1} = \mathbf{1}_A + \sum_{n\in\mathbb{N}} \lambda^n x^n.$$

This series is called the **Neumann series for** x.

Corollary 12.10. *Let* $(A, \|\cdot\|)$ *be a unital Banach algebra. Then* $B(\mathbf{1}_A, 1) \subseteq A^{-1}$.

Proof. The proof is left as an exercise for the reader. □

Corollary 12.11. *Let* $(A, \|\cdot\|)$ *be a unital Banach algebra. If* $x \in A$ *and* $\|x\| < 1$*, then*

$$\|(\mathbf{1}_A - x)^{-1}\| \leqslant \frac{1}{1 - \|x\|}.$$

Proof. From Theorem 12.8,

$$\|(\mathbf{1}_A - x)^{-1}\| = \left\| \mathbf{1}_A + \sum_{k=1}^{\infty} x^k \right\| \leqslant 1 + \sum_{k=1}^{\infty} \|x^k\| \leqslant \sum_{k=0}^{\infty} \|x\|^k = \frac{1}{1 - \|x\|}.$$

This completes the proof. □

Corollary 12.12. *Let* $(A, \|\cdot\|)$ *be a unital Banach algebra, then* A^{-1} *is an open set.*

Proof. Let $x_0 \in A^{-1}$. Then $x_0 \in x_0 \cdot B(\mathbf{1}_A, 1) \subseteq A^{-1}$, since $B(\mathbf{1}_A, 1) \subseteq A^{-1}$. Now, $x_0 \cdot B(\mathbf{1}_A, 1)$ is open in $(A, \|\cdot\|)$ and so $x_0 \in \text{int}(A^{-1})$; which completes the proof. □

Theorem 12.13. *Let* $(A, \|\cdot\|)$ *be a unital Banach algebra, then* $x \mapsto x^{-1}$ *is continuous on* A^{-1}*. In fact,* $(A^{-1}, \cdot)$ *is a topological group.*

Proof. Suppose $x, y \in A^{-1}$, then

$$\|y^{-1} - x^{-1}\| \quad = \quad \|x^{-1}(x - y)y^{-1}\| \leqslant \|x^{-1}\| \cdot \|(x - y)\| \cdot \|y^{-1}\|$$

and since $x^{-1} = y^{-1} + (x^{-1} - y^{-1})$

$$\|x^{-1}\| \quad \leqslant \quad \|y^{-1}\| + \|y^{-1} - x^{-1}\| \leqslant \|y^{-1}\| + \|x^{-1}\| \cdot \|x - y\| \cdot \|y^{-1}\|.$$

Note that this immediately implies that

$$\|x^{-1}\| \cdot \big(1 - \|x - y\| \cdot \|y^{-1}\|\big) \leqslant \|y^{-1}\|$$

or

$$\|x^{-1}\| \leqslant \frac{\|y^{-1}\|}{(1 - \|x - y\| \cdot \|y^{-1}\|)}$$

provided $\|x - y\| < 1/\|y^{-1}\|$. This then gives us that

$$\|x^{-1} - y^{-1}\| \leqslant \frac{\|x - y\| \cdot \|y^{-1}\|^2}{(1 - \|x - y\| \cdot \|y^{-1}\|)} \leqslant 2\|x - y\| \cdot \|y^{-1}\|^2$$

provided $0 \leqslant \|x - y\| < 1/2\|y^{-1}\|$. Thus, given $\varepsilon > 0$, if we choose

$$\delta := \min \left\{ \frac{1}{2\|y^{-1}\|}, \frac{\varepsilon}{2\|y^{-1}\|^2} \right\} > 0$$

then $\|x^{-1} - y^{-1}\| < \varepsilon$ whenever $\|x - y\| < \delta$. $\square$

Unitisation

Theorem 12.14. *If $(A, \|\cdot\|)$ is a Banach algebra without an identity element, then there exists a unital Banach algebra $(B, |\!|\!| \cdot |\!|\!|)$ such that A is a closed subalgebra of B.*

Proof. Let $B := A \times \mathbb{C}$ and define,

$$(x, a) + (y, b) := (x + y, a + b), (x, a)(y, b) := (xy + ay + bx, ab), \lambda(x, a) := (\lambda x, \lambda a).$$

Also define $|\!|\!|(x, a)|\!|\!| := \|x\| + |a|$. Then $(B, |\!|\!| \cdot |\!|\!|)$ is a Banach algebra with identity $\mathbf{1}_B := (0, 1)$ and A is isometrically isomorphic to $A \times \{0\}$. $\square$

Application

Suppose that $f, g \in C_{\mathbb{C}}[a, b]$ and that k is a continuous complex-valued function defined on the triangular region $\{(x, t) \in [a, b] \times [a, b] : a \leqslant t \leqslant x\}$. Then the **Volterra integral equation** determined by f, g, k and $\lambda \in \mathbb{C}$ is the equation:

$$f(x) = g(x) + \lambda \int_{[a,x]} k(x, t) f(t) \, \mathrm{d}t \quad \text{for all } x \in [a, b].$$

Theorem 12.15. *For each $g \in C_{\mathbb{C}}[a, b]$ and continuous complex-valued function k defined on the triangular region $\{(x, t) \in [a, b] \times [a, b] : a \leqslant t \leqslant x\}$. The Volterra equation*

$$f(x) = g(x) + \lambda \int_{[a,x]} k(x, t) f(t) \, \mathrm{d}t \quad \textit{for all } x \in [a, b]$$

has a unique solution for every $\lambda \in \mathbb{C}$.

Proof. We define the **Volterra operator** $K : (C_{\mathbb{C}}[a,b], \|\cdot\|_\infty) \to (C_{\mathbb{C}}[a,b], \|\cdot\|_\infty)$ by,

$$K(f)(x) := \int_{[a,x]} k(x,t)f(t)\ \mathrm{d}t.$$

It is a straightforward exercise (which we leave to the reader) to show that K is a continuous linear operator on $C_{\mathbb{C}}[a,b]$. In terms of the Volterra operator, the Volterra integral equation can be written as $(I - \lambda K)(f) = g$. From before, we see that $(I - \lambda K)$ is invertible (i.e., regular) for all $\lambda \in \mathbb{C}$, provided that $\lim_{n\to\infty} \|K^n\|^{\frac{1}{n}} = 0$ and furthermore the solution will be given by the Neumann series

$$f = (I - \lambda K)^{-1}(g) = \big(I + \textstyle\sum_{n\in\mathbb{N}} \lambda^n K^n\big)(g).$$

That is, we have a series solution for the Volterra integral equation. So next we will show that $\lim_{n\to\infty} \|K^n\|^{\frac{1}{n}} = 0$. Now,

$$\begin{aligned} |K(f)(x)| &\leqslant \int_{[a,x]} |k(x,t)||f(t)|\ \mathrm{d}t \\ &\leqslant (x-a)\sup\{|k(x,t)||f(t)| : a \leqslant t \leqslant x\} \\ &\leqslant M\|f\|_\infty (x-a) \end{aligned}$$

where $M := \sup\{|k(x,t)| : a \leqslant t \leqslant x \text{ and } a \leqslant x \leqslant b\}$. We shall prove by induction that

$$|K^n(f)(x)| \leqslant M^n \|f\|_\infty \frac{(x-a)^n}{n!} \quad \text{for all } a \leqslant x \leqslant b.$$

We have already shown that this is true in the case when $n = 1$. So suppose that the statement is true for the case $n = m$. Then,

$$\begin{aligned} |K^{m+1}(f)(x)| = |K(K^m(f))(x)| &= \left|\int_{[a,x]} k(x,t)(K^m(f))(t)\ \mathrm{d}t\right| \\ &\leqslant \int_{[a,x]} |k(x,t)||(K^m(f))(t)|\ \mathrm{d}t \\ &\leqslant \frac{M^m\|f\|_\infty}{m!}\int_{[a,x]} M(t-a)^m\ \mathrm{d}t \\ &\leqslant \frac{M^{m+1}\|f\|_\infty (x-a)^{m+1}}{(m+1)!} \end{aligned}$$

which concludes the induction. Using this fact we obtain that for all $n \in \mathbb{N}$,

$$\|K^n(f)\|_\infty = \max\{|K^n(f)(x)| : a \leqslant x \leqslant b\} \leqslant M^n \|f\|_\infty \frac{(b-a)^n}{n!}$$

and so

$$\|K^n\| = \sup\{\|K^n(f)\|_\infty : \|f\|_\infty \leqslant 1\} \leqslant M^n \frac{(b-a)^n}{n!}.$$

Since $\lim_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = 0$ we conclude that $\lim_{n\to\infty} \|K^n\|^{\frac{1}{n}} = 0$. $\square$

Chapter 13

The Resolvent Function

Let $(A, \|\cdot\|)$ be a unital Banach algebra. We define the **spectrum of** $x \in A$ to be

$$\sigma_A(x) := \{\lambda \in \mathbb{C} : x - \lambda \mathbf{1}_A \text{ is singular}\}.$$

When there is no ambiguity we shall simply write $\sigma(x)$ for $\sigma_A(x)$. Recall that an element $a \in A$ is called **singular** if $a \notin A^{-1}$.

It is easy to see that $\lambda \mapsto (x - \lambda \mathbf{1}_A)$ is a continuous function from $\mathbb{C}$ into A. Since the set of singular elements in A is closed, it follows at once that $\sigma_A(x)$ is closed. Further, observe that $\sigma_A(x) \subseteq \{z \in \mathbb{C} : |z| \leqslant \|x\|\}$ because if $\lambda > \|x\|$, then $(\mathbf{1}_A - \lambda^{-1}x)$ is a unit, since $\left\|\lambda^{-1}x\right\| < 1$ and so $(x - \lambda\mathbf{1}_A)$ is a unit as well, since $(x - \lambda\mathbf{1}_A) = (-\lambda)(\mathbf{1}_A - \lambda^{-1}x)$. Thus, for each $x \in A$, $\sigma_A(x)$ is compact.

Basic facts: Let $(A, \|\cdot\|)$ be a unital Banach algebra.

(i) If A is a subalgebra of a Banach algebra $(B, \|\cdot\|)$, then $\sigma_B(x) \subseteq \sigma_A(x)$ for all $x \in A$.

(ii) If $\lambda \in \mathbb{C}$ and $x \in A$, then $\sigma_A(\lambda x) = \lambda\sigma_A(x)$.

(iii) If $\lambda \in \mathbb{C}$ and $x \in A$, then $\sigma_A(x + \lambda\mathbf{1}_A) = \sigma_A(x) + \lambda$.

(iv) If B is a Banach algebra and $\pi : A \to B$ is a unital homomorphism (i.e., an algebra homomorphism such that $\pi(\mathbf{1}_A) = \mathbf{1}_B$), then $\sigma_B(\pi(x)) \subseteq \sigma_A(x)$.

(v) If $x \in A^{-1}$ and $\lambda \in \mathbb{C} \setminus \{0\}$, then $(x^{-1} - \lambda^{-1}\mathbf{1}_A) = (-\lambda)^{-1}x^{-1}(x - \lambda\mathbf{1}_A)$.

(vi) If $x \in A^{-1}$, then $\sigma_A(x^{-1}) = \{\lambda^{-1} : \lambda \in \sigma_A(x)\}$.

(vii) If $x, y \in A$ and $(\mathbf{1}_A - xy) \in A^{-1}$, then $(\mathbf{1}_A - yx) \in A^{-1}$. *Hint*: Consider the element $\mathbf{1}_A + y(\mathbf{1}_A - xy)^{-1}x$.

(viii) For any $x, y \in A$, $\sigma_A(xy) \setminus \{0\} = \sigma_A(yx) \setminus \{0\}$.

Proof. We give only outlines.

(i) This follows from the fact that $A^{-1} \subseteq B^{-1}$.

(ii) Check first that $\sigma_A(0x) = 0\sigma(x) = \{0\}$, assuming we know that $\sigma_A(x) \neq \varnothing$. Then check that $\sigma_A(\lambda x) = \lambda\sigma_A(x)$ for $\lambda \neq 0$.

(iii) Straightforward.

(iv) Firstly note that $\pi(A^{-1}) \subseteq B^{-1}$. Indeed, if $\mathbf{1}_A = ab = ba$, then

$$\mathbf{1}_B = \pi(\mathbf{1}_A) = \pi(ab) = \pi(a)\pi(b) \quad \text{and} \quad \mathbf{1}_B = \pi(\mathbf{1}_A) = \pi(ba) = \pi(b)\pi(a).$$

Therefore, $\pi(a) \in B^{-1}$. Now, suppose that $\lambda \notin \sigma_A(x)$, then $(x - \lambda\mathbf{1}_A) \in A^{-1}$ and so $\pi(x) - \lambda\mathbf{1}_B = \pi(x - \lambda\mathbf{1}_A) \in B^{-1}$, i.e., $\lambda \notin \sigma_B(\pi(x))$.

(v) Straightforward.

(vi) Again straightforward.

(vii) To check this, one just does the multiplication, but to see where this formula might have come from, consider the following formal calculation

$$(\mathbf{1}_A - xy)^{-1} = \mathbf{1}_A + xy + (xy)^2 + \cdots = \mathbf{1}_A + x(\mathbf{1}_A + yx + (yx)^2 + \cdots)y = \mathbf{1}_A + x(\mathbf{1}_A - yx)^{-1}y.$$

(viii) This just follows from (vii).

This completes the justifications of the basic facts. □

Example 13.1. *We consider some basic examples.*

(i) Let $A := M_n(\mathbb{C})$ and define $\|M\| := \sup\{\|M\boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\}$. Then $(A, \|\cdot\|)$ is a unital Banach algebra and for each $M \in A$, $\sigma_A(M)$ consists of all the eigenvalues of M.

(ii) Let $A = C_{\mathbb{C}}(K)$, then for each $f \in A$, $\sigma_A(f) = \{f(k) : k \in K\}$, i.e., $\sigma_A(f)$ is the image of f. To see this, note that if $f \in A$, then

$$\begin{aligned} \lambda \notin \sigma_A(f) &\iff (f - \lambda\mathbf{1}) \textit{ is invertible} \\ &\iff (f - \lambda\mathbf{1})(x) \neq 0 \textit{ for any } x \in K \\ &\iff \lambda \neq f(x) \textit{ for any } x \in K \\ &\iff \lambda \notin f(K). \end{aligned}$$

(iii) Let H be a Hilbert space. For $T \in B(H)$, $\sigma_{B(H)}(T)$ contains all the eigenvalues, but could be strictly larger. For example, take $H = \ell^2(\mathbb{N})$ and let T be defined by,

$$T(x_1, x_2, x_3, \ldots) := (0, x_1, x_2, x_3, \ldots).$$

We claim that (a) T has no eigenvalues and (b) $\sigma_{B(H)}(T) = \{\lambda \in \mathbb{C} : |\lambda| \leqslant 1\}$. To prove (a) suppose λ is an eigenvalue so that there exists a nonzero sequence $(x_n : n \in \mathbb{N}) \in \ell^2(\mathbb{N})$ with $T[(x_n : n \in \mathbb{N})] = \lambda(x_n : n \in \mathbb{N})$. Then

$$(0, x_1, x_2, x_3, \ldots) = (\lambda x_1, \lambda x_2, \lambda x_3, \ldots);$$

the left-hand side is nonzero, so λ cannot be zero. Also it follows that $\lambda x_1 = 0$, i.e., $x_1 = 0$ and $x_n = \lambda x_{n+1}$ for all $n \in \mathbb{N}$, i.e., $x_{n+1} = \lambda^{-n} x_1$ for all $n \in \mathbb{N}$. Therefore, $x_n = 0$ for all $n \in \mathbb{N}$, which is a contradiction.

(b) Since $\|T\| = 1$ we know from above that $\sigma_{B(H)}(T) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leqslant 1\}$. So let us show that if $|\lambda| \leqslant 1$, then $T - \lambda\mathbf{1}$ is not surjective by showing that $(1, 0, 0, 0, \ldots)$ is not in the

range of $(T-\lambda\mathbf{1})$*. Suppose that* $(x_n : n \in \mathbb{N}) \in \ell^2(\mathbb{N})$ *satisfies* $(T-\lambda\mathbf{1})[(x_n : n \in \mathbb{N})] = (1,0,0,0,\ldots)$*. Then*

$$(0-\lambda x_1, x_1 - \lambda x_2, x_2 - \lambda x_3, \ldots) = (1,0,0,0,\ldots).$$

Since $-\lambda x_1 = 1$*,* $x_1 = -1/\lambda$*. Moreover, since* $x_n - \lambda x_{n+1} = 0$ *for all* $n \in \mathbb{N}$ *we have that* $x_{n+1} = \lambda^{-n} x_1$ *for all* $n \in \mathbb{N}$*, i.e.,* $x_{n+1} = -\lambda^{-(n+1)}$ *for all* $n \in \mathbb{N}$*, but then* $(x_n : n \in \mathbb{N}) \notin \ell^2(\mathbb{N})$*. This gives (b).* $\square$

Proposition 13.2. *Suppose that* $(B, \|\cdot\|)$ *is a unital Banach algebra and* $(A, \|\cdot\|)$ *is a Banach subalgebra of* B*, with* $\mathbf{1}_B \in A$*. Then for any* $x \in A$*,* $\partial\sigma_A(x) \subseteq \sigma_B(x) \subseteq \sigma_A(x)$*. Here,* $\partial\sigma_A(x)$ *denotes the boundary of* $\sigma_A(x)$*.*

Proof. As $A^{-1} \subseteq B^{-1}$ it follows that $\sigma_B(x) \subseteq \sigma_A(x)$. So we consider the other set inclusion. To obtain a contradiction, let us suppose there is some $\lambda \in \partial\sigma_A(x) \setminus \sigma_B(x)$. Then $(x-\lambda\mathbf{1}_B)^{-1} \in B \setminus A$. Since $\lambda \in \partial\sigma_A(x)$ there exists a sequence $(\lambda_n : n \in \mathbb{N})$ in $\mathbb{C} \setminus \sigma_A(x)$ such that $\lambda = \lim_{n\to\infty} \lambda_n$. Therefore, $(x-\lambda_n\mathbf{1}_B)^{-1} \in A$ for all $n \in \mathbb{N}$ and

$$(x-\lambda\mathbf{1}_B)^{-1} = \left(\lim_{n\to\infty}(x-\lambda_n\mathbf{1}_B)\right)^{-1} = \lim_{n\to\infty}(x-\lambda_n\mathbf{1}_B)^{-1} \in A$$

since A is closed and the mapping $b \mapsto b^{-1}$ is continuous on B^{-1}. However, this contradicts the assumption that $\lambda \in \sigma_A(x)$. $\square$

Example 13.3. *Let* $D := \{x \in \mathbb{C} : |z| \leqslant 1\}$ *and let* $\mathbb{T} := \partial D$*, i.e.,* $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$*. Let* $A(D) := \{f \in C_{\mathbb{C}}(D) : f$ *is analytic on* $\mathrm{int}(D)\}$*. Then* $(A(D), \|\cdot\|_\infty)$ *is a unital Banach algebra and* $\sigma_{A(D)}(f) = f(D)$ *for every* $f \in A(D)$*.*

Proof. From Example 12.2 part(ii) we know that $(A(D), \|\cdot\|_\infty)$ is a unital Banach algebra and by Exercise 12.5 part(ii) we know that $f \in A(D)$ is invertible if, and only if, $0 \notin f(D)$. From this it follows that $\sigma_{A(D)}(f) = f(D)$. Let $R : A(D) \to C(\mathbb{T})$ be defined by, $R(f) := f|_{\mathbb{T}}$. Then by the Maximum Modulus Principle, $\|R(f)\|_\infty = \|f\|_\infty$ for all $f \in A(D)$. Therefore, R is a Banach algebra isomorphism from $A(D)$ onto $R(A(D))$. Let $X := A(D)$ and $Y := R(A(D))$. Then $\sigma_X(f) = \sigma_Y(R(f))$ for all $f \in X$. In particular, $\sigma_Y(R(id_D)) = \sigma_X(id_D) = D$, where $id_D : D \to \mathbb{C}$ is defined by, $id_D(z) := z$ for all $z \in D$. Let $g := R(id_D))$, then $\sigma_Y(g) = D$.

On the other hand, Y is a subalgebra of $Z := C_{\mathbb{C}}(\mathbb{T})$ and $\sigma_Z(g) = \mathbb{T}$. Thus,

$$\sigma_Z(g) = \mathbb{T} = \partial D = \partial[\sigma_Y(g)].$$

This completes the exmple. $\square$

Let $(A, \|\cdot\|)$ be a unital Banach algebra, then the **resolvent of** $x \in A$ is the function $R : \mathbb{C} \setminus \sigma_A(x) \to A$ defined by,

$$R(\lambda) := (x-\lambda\mathbf{1})^{-1}.$$

Since $R(\lambda) = (-\lambda)^{-1}(\mathbf{1}-\lambda^{-1}x)^{-1}$ for $\lambda \in \mathbb{C}\setminus\sigma_A(x)$ we have that $\|R(\lambda)\| \to 0$ as $|\lambda| \to \infty$.

If μ and $\lambda \in \mathbb{C} \setminus \sigma_A(x)$, then

$$R(\mu) - R(\lambda) = R(\lambda)(\mu\mathbf{1} - \lambda\mathbf{1})R(\mu) = (\mu - \lambda)R(\lambda)R(\mu).$$

Thus, if $x^* \in A^*$, then

$$\frac{x^*(R(\mu)) - x^*(R(\lambda))}{\mu - \lambda} = x^*(R(\lambda)R(\mu))$$

for all $\mu, \lambda \in \mathbb{C} \setminus \sigma(x)$ with $\mu \neq \lambda$.

The next theorem requires a result from complex analysis, namely Liouville's Theorem, which says that the only bounded analytic functions $f : \mathbb{C} \to \mathbb{C}$ are the constant functions.

Theorem 13.4. *Let $(A, \|\cdot\|)$ be a unital Banach algebra and let $a \in A$. Then $\sigma_A(a) \neq \varnothing$.*

Proof. Fix $x^* \in A^*$ and define $f : \mathbb{C} \setminus \sigma_A(x) \to \mathbb{C}$ by, $f(\lambda) := x^*(R(\lambda))$. Then for any $\lambda, \mu \in \mathbb{C}\backslash\sigma_A(x)$, $(\lambda \neq \mu)$,

$$\frac{f(\mu) - f(\lambda)}{\mu - \lambda} = x^*(R(\lambda)R(\mu))$$

Thus,

$$f'(\lambda) = \lim_{\mu \to \lambda} \frac{f(\mu) - f(\lambda)}{\mu - \lambda} = x^*(R^2(\lambda)), \text{ since } R \text{ is continuous on } \mathbb{C}\backslash\sigma_A(x).$$

So f is analytic on $\mathbb{C} \setminus \sigma_A(x)$. Moreover, for any $\lambda \in \mathbb{C}\backslash\sigma_A(x)$,

$$|f(\lambda)| \leqslant \|x^*\| \, \|R(\lambda)\| = (\|x^*\|/|\lambda|)\| \left(\mathbf{1} - \lambda^{-1}x\right)^{-1} \|.$$

Therefore $|f(\lambda)| \to 0$ as $|\lambda| \to \infty$.

Now suppose, in order to obtain a contradiction, that $\sigma_A(x) = \varnothing$. Then f is a bounded entire function (i.e., analytic on all of $\mathbb{C}$) and so from Liouville's Theorem $f \equiv c$ for some $c \in \mathbb{C}$. However, since $f \to 0$ as $|\lambda| \to \infty$ it must be the case that $f \equiv 0$. Therefore, for each $x^* \in A$, $x^*(R(\lambda)) = 0$ for all $\lambda \in \mathbb{C}$. Hence, by the Hahn-Banach Theorem $R(\lambda) = 0$ for all $\lambda \in \mathbb{C}$. However, this is absurd since 0 is not invertible. □

An algebra with identity in which each nonzero element is invertible is called a **division algebra**.

Theorem 13.5 (Gelfand-Mazur)**.** *If $(A, \|\cdot\|)$ is a division Banach algebra, then it equals the set of all scalar multiples of the identity.*

Proof. Let $x \in A$ and $\lambda \in \sigma_A(x) \neq \varnothing$. Then $x - \lambda\mathbf{1}_A$ must equal 0, i.e., $x = \lambda\mathbf{1}_A$. □

For an element x of a unital Banach algebra $(A, \|\cdot\|)$ we define the **spectral radius of** x to be

$$r_A(x) := \max\{|\lambda| : \lambda \in \sigma_A(x)\}.$$

When there is no ambiguity we simply write $r(x)$ for $r_A(x)$.

We need some further results from complex analysis. Recall that if f is analytic in a ball $B(z_0, r)$, then the Taylor series for f converges to f throughout $B(z_0, r)$. We need the following analogue for functions analytic in an annulus

$$A(z_0, r, R) := \{z \in \mathbb{C} : r < |z| < R\}.$$

Theorem 13.6. *Suppose that the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ converges when $|z-z_0| < R$ and $\sum_{n=-\infty}^{-1} a_n(z-z_0)^n$ converges when $|z-z_0| > r$. Then the function $f : A(z_0, r, R) \to \mathbb{C}$ defined by the following Laurent series*

$$f(x) = \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n := \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n + \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$$

is analytic in $A(z_0, r, R)$. Conversely, if $f : A(z_0, r, R) \to \mathbb{C}$ is analytic, then there is a unique Laurent series which converges absolutely to $f(z)$ for every $z \in A(z_0, r, R)$.

Theorem 13.7 (Spectral Radius Formula)**.** *Let $(A, \|\cdot\|)$ be a unital Banach algebra and let $x \in A$. Then*

$$r_A(x) = \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}.$$

Proof. Note that $r_A(x) \leqslant \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}$ since if $\lambda > \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}$, then $\lim_{n\to\infty} \left\|\left(\lambda^{-1}x\right)^n\right\|^{\frac{1}{n}} < 1$ and so $\left(\mathbf{1} - \lambda^{-1}x\right)$ is a unit. However,

$$(x - \lambda\mathbf{1}) = (-\lambda)\left(\mathbf{1} - \lambda^{-1}x\right)$$

and so $(x - \lambda\mathbf{1})$ is a unit as well, i.e., $\lambda \notin \sigma_A(x)$.

So now we need only show that $r_A(x) \geqslant \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}$. To do this, it suffices to show that if $r_A(x) < a$, then $\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} \leqslant a$.

For $|\lambda| > \|x\|$ we have that

$$R(\lambda) = (x - \lambda\mathbf{1})^{-1} = (-\lambda)^{-1}\left(\mathbf{1} - \lambda^{-1}x\right)^{-1} = (-\lambda)^{-1}\sum_{k=0}^{\infty} \lambda^{-k}x^k.$$

Fix $x^* \in A^*$ and define $f : \mathbb{C} \setminus \sigma_A(x) \to \mathbb{C}$ by, $f(\lambda) := x^*(R(\lambda))$. For $|\lambda| > \|x\|$, and in particular, for $\lambda \in A(0, \|x\|, \|x\| + 1)$ we have that

$$f(\lambda) = (-\lambda)^{-1}\sum_{k=0}^{\infty} \frac{x^*(x^k)}{\lambda^k}.$$

As we have seen previously, f is analytic on $\mathbb{C} \setminus \sigma_A(x)$. Therefore, f has a Laurent expansion on $A(0, r_A(x), \|x\| + 1)$. Moreover, since the Laurent expansion of f is unique it must coincide with the Laurent expansion given above on the annulus $A(0, \|x\|, \|x\|+1)$. Hence,

$$f(\lambda) = (-\lambda)^{-1} \sum_{k=0}^{\infty} \frac{x^*(x^k)}{\lambda^k} \quad \text{for } \lambda \in A(0, r_A(x), \|x\| + 1).$$

Therefore,

$$f(a) = (-a)^{-1} \sum_{k=0}^{\infty} \frac{x^*(x^k)}{a^k} = (-a)^{-1} \sum_{k=0}^{\infty} x^* \left(a^{-k} x^k\right).$$

In particular, $\lim_{n\to\infty} x^* (a^{-n} x^n) = 0$ and so the set $\{x^*(a^{-n}x^n) : n \in \mathbb{N}\}$ is bounded. Since this holds for any $x^* \in A^*$ the set $\{a^{-n}x^n : n \in \mathbb{N}\}$ is weakly bounded and hence, by the Uniform Boundedness Theorem, norm bounded. That is, there exists a $K > 0$ such that $\|x^n\| \leqslant K a^n$ for all $n \in \mathbb{N}$. Thus, $\|x^n\|^{\frac{1}{n}} \leqslant K^{\frac{1}{n}} a$ for all $n \in \mathbb{N}$ and so $\lim\limits_{n\to\infty} \|x^n\|^{\frac{1}{n}} \leqslant 1a = a$. $\square$

Let A be an algebra. Then a linear functional x^* on A is called a **multiplicative linear functional** if $x^*(xy) = x^*(x)x^*(y)$ for all $x, y \in A$.

Note that if K is a compact topological space and $x \in K$, then $\delta_x : C_{\mathbb{C}}(K) \to \mathbb{C}$ defined by, $\delta_x(f) := f(x)$ for all $f \in C_{\mathbb{C}}(K)$ is a multiplicative linear functional on $C_{\mathbb{C}}(K)$.

Remarks 13.8. *Let A be an algebra. Then $x^* : A \to \mathbb{C}$ is a multiplicative linear functional on A if, and only if, x^* is an algebra homomorphism.*

Exercise 13.9. *These exercises are on multiplicative linear functionals.*

(i) Show that if A is an algebra with identity and x^* is a nonzero multiplicative linear functional on A, then $x^*(\mathbf{1}_A) = 1$.

(ii) Show that if $(A, \|\cdot\|)$ is a Banach algebra and $x^* \in A^*$ is a multiplicative linear functional on $(A, \|\cdot\|)$, then $\|x^*\| \leqslant 1$. Hint: Suppose to the contrary that there exists an element $x' \in B_A$ such that $|x^*(x')| > 1$. Then show that this implies that there exists an element $x \in A$ such that $\|x\| < 1$ and $x^*(x) = 1$. Let $x := \frac{1}{x^*(x')} x'$ and consider $y := \sum_{n=1}^{\infty} x^n$ and show that $x + xy = y$. Finally, deduce that this leads to a contradiction.

(iii) Show that if $(A, \|\cdot\|)$ is a unital Banach algebra and $x^* \in A^*$ is a nonzero multiplicative linear functional on $(A, \|\cdot\|)$, then $\|x^*\| = 1$.

Let $(A, \|\cdot\|)$ be a unital Banach algebra. We call a functional $x^* \in A^*$ a **state** if $\|x^*\| = x^*(\mathbf{1}) = 1$. We shall denote by $S(A)$ the set of all state functionals in A^* and by Δ_A the set of all nonzero multiplicative linear functionals on A. We know from the previous exercises that $\Delta_A \subseteq S(A) \subseteq S_{A^*}$.

Note that if K is a compact Hausdorff topological space and p is a Borel probability measure on K, then $x^* : C_{\mathbb{C}}(K) \to \mathbb{C}$ defined by, $x^*(f) := \int_K f \, \mathrm{d}p$ for all $f \in C_{\mathbb{C}}(K)$, is a state on $C_{\mathbb{C}}(K)$.

Recall that a subset U in the dual of a normed linear space $(X, \|\cdot\|)$ is called **weak* open** if for each $x^* \in U$ there exists an $\varepsilon > 0$ and a finite set $\{x_1, x_2, \ldots x_n\}$ in X such that the set

$$N(x^*, x_1, x_2, \ldots x_n, \varepsilon) := \{y^* \in X^* : |x^*(x_j) - y^*(x_j)| < \varepsilon \text{ for each } 1 \leqslant j \leqslant n\}$$

is contained in U.

Exercise 13.10. *Let $(X, \|\cdot\|)$ be a normed linear space.*

(i) Show that the set of all weak* open subsets of X^* forms a topology on X^*. This topology is called the **weak* topology** *on X^*.*

(ii) Show that the weak* topology on X^* is weaker than the norm topology on X^*.

(iii) Show that each element of $\widehat{X}$ is continuous on $(X^*, weak^*)$.

Let $(X, \|\cdot\|)$ be a normed linear space. Then the weak* topology on X^* is sometimes called the topology of **pointwise convergence on** X. Furthermore, it can be shown that the weak* topology on X^* is the weakest topology on X^* that make each functional from $\widehat{X}$ continuous, i.e., the weak* topology on X^* is the *weak topology* on X^* generated by $\widehat{X}$.

Theorem 13.11 (Banach-Alaoglu Theorem)**.** *Let $(X, \|\cdot\|)$ be a normed linear space, then $(B_{X^*}, weak^*)$ is compact.*

Exercise 13.12. *Let $(A, \|\cdot\|)$ be a unital Banach algebra. Show that $S(A)$ is a weak* compact convex subset of A^*. Hint: $S(A) = B_{A^*} \cap (\widehat{\mathbf{1}_A})^{-1}(1)$.*

Theorem 13.13. *Let $(A, \|\cdot\|)$ be a unital Banach algebra. Then Δ_A is a weak* closed and hence a weak* compact subset of B_{A^*}.*

Proof. Firstly, as already noted, $\Delta_A \subseteq S(A) \subseteq B_A$. So it is sufficient to show that Δ_A is weak* closed.

$$\begin{aligned}
\Delta_A &= \textstyle\bigcap_{x,y\in A}\{x^* \in A^* : x^*(\mathbf{1}_A) = 1 \text{ and } x^*(xy) = x^*(x)x^*(y)\} \\
&= \textstyle\bigcap_{x,y\in A}\{x^* \in A^* : x^*(\mathbf{1}_A) = 1 \text{ and } (\widehat{xy} - \widehat{x}\widehat{y})(x^*) = 0\} \\
&= (\widehat{\mathbf{1}_A})^{-1}(1) \cap \textstyle\bigcap_{x,y\in A} q_{x,y}^{-1}(0), \quad \text{where } q_{x,y} := \widehat{xy} - \widehat{x}\widehat{y}.
\end{aligned}$$

Since each $q_{x,y}$ is weak* continuous, $q_{x,y}^{-1}(0)$ is weak* closed. Therefore, Δ_A being the intersection of weak* closed subsets is itself weak* closed. $\square$

Eventually, we will show that if $(A, \|\cdot\|)$ is a commutative unital Banach algebra, then there exists an algebra homomorphism $\varphi : (A, \|\cdot\|) \to (C_{\mathbb{C}}(\Delta_A), \|\cdot\|_\infty)$ such that for each $x \in A$, $\|\varphi(x)\|_\infty = r(x)$.

To prove this we first need to prove three preliminary results.

Let A be an algebra, then a subset I of A is called a **2-sided ideal** if:

(i) I is a vector subspace of A;

(ii) $xI \subseteq I$ and $Ix \subseteq I$ for all $x \in A$.

Using Zorn's Lemma it is easy to show that every proper ideal in a unital algebra is contained in a maximal, with respect to set inclusion, proper ideal.

If A is a commutative algebra with identity and $x \in A$, then the set $\{ax : a \in A\}$ is an ideal in A and is called the **principal ideal generated by** x and is denoted by $\langle x \rangle$. An ideal I is called a **principal ideal** if $I = \langle x \rangle$ for some $x \in I$.

Lemma 13.14. *Let A be a commutative algebra with identity. Then every singular element $x \in A$ is contained in a maximal proper ideal. In fact, $x \in A$ is singular if, and only if, it is contained in a maximal proper ideal.*

Proof. If x is singular, then $\langle x \rangle$ is a proper ideal in A, since $\mathbf{1}_A \notin \langle x \rangle$. Hence, by the above, there exists a maximal proper ideal N such that $x \in \langle x \rangle \subseteq N$. Conversely, if x is nonsingular (i.e., invertible) and N is an ideal in A containing x, then $N = A$. Thus, if x is a unit in A, then x is not contained in any proper ideal in A. $\square$

Note: If $(A, \|\cdot\|)$ is a unital Banach algebra, then each maximal ideal is closed, since if I is an ideal in A, then so is $\overline{I}$. Moreover, if I is a proper ideal in A, then $I \cap A^{-1} = \varnothing$. Therefore, $\overline{I} \cap A^{-1} = \varnothing$ and so $\overline{I}$ is also a proper ideal in A.

Lemma 13.15. *If I is a proper closed 2-sided ideal in a Banach algebra $(A, \|\cdot\|)$. Then the quotient Banach space A/I is a Banach algebra in which $(a+I)(b+I) = (ab+I)$. The quotient map $q : a \mapsto a + I$ is a norm-decreasing homomorphism with kernel I. Furthermore, if $(A, \|\cdot\|)$ is a unital Banach algebra, then so is A/I, with multiplicative identity $\mathbf{1}_A + I$.*

Proof. It is routine to check that if I is an ideal, then $(a+I)(b+I) = (ab+I)$ gives a well-defined multiplication on A/I. Indeed, if $a + I = a' + I$ and $b + I = b' + I$, then $a = a' + x$ and $b = b' + y$ for some $x, y \in I$ and

$$ab = (a' + x)(b' + y) = a'b' + (a'y + xb' + xy);$$

because I is a 2-sided ideal, $a'y + xb' + xy \in I$ and so $ab + I = a'b' + I$. Associativity, distributivity and the properties of the identity $\mathbf{1}_A + I$ all follow immediately from the corresponding properties of A. We know from our work on Banach spaces that if I is closed in $(A, \|\cdot\|)$ then A/I is a Banach space in the quotient norm

$$\|a + I\| := \inf\{\|a + x\| : x \in I\}.$$

To see that the norm is submultiplicative, let $a+I$, $b+I \in A/I$. Then,

$$\begin{aligned}
\|(a+I)(b+I)\| &= \|ab+I\| \\
&= \inf_{w\in I}\|ab+w\| \\
&\leqslant \inf_{z,z'\in I}\|ab+(az'+zb+zz')\| \\
&= \inf_{z,z'\in I}\|(a+z)(b+z')\| \\
&\leqslant \inf_{z,z'\in I}\|a+z\|\|b+z'\| \\
&= \left(\inf_{z\in I}\|a+z\|\right)\left(\inf_{z'\in I}\|b+z'\|\right) \\
&= \|a+I\|\|b+I\|
\end{aligned}$$

i.e., $\|(a+I)(b+I)\| \leqslant \|a+I\|\|b+I\|$; which shows that the norm on A/I is submultiplicative. In particular,

$$\|\mathbf{1}_A+I\| = \|(\mathbf{1}_A+I)(\mathbf{1}_A+I)\| \leqslant \|\mathbf{1}_A+I\|^2.$$

Since $\mathbf{1}_A \notin I$, $\mathbf{1}_A+I \neq I$ and so $0 < \|\mathbf{1}_A+I\|$. Therefore, $1 \leqslant \|\mathbf{1}_A+I\|$. On the other hand, $\|\mathbf{1}_A+I\| = \inf\{\|\mathbf{1}_A+x\| : x \in I\} \leqslant \|\mathbf{1}_A+0\| = \|\mathbf{1}_A\| = 1$, since $0 \in I$. Thus, $\|\mathbf{1}_A+I\| = 1$, which shows that A/I is a Banach algebra. That q is norm decreasing follows from the definition of the quotient norm, that q is a homomorphism follows from the definition of scalar multiplication, addition and multiplication in A/I. $\square$

Lemma 13.16. *Let N be a maximal proper ideal in a commutative unital Banach algebra $(A, \|\cdot\|)$. Then there exists a nonzero multiplicative linear functional x^* on A such that $N = \mathrm{Ker}(x^*)$.*

Proof. Firstly, from our earlier note, we know that N is closed. Therefore, by Lemma 13.15 we know that A/N is a unital Banach algebra. We claim that A/N is a division Banach algebra. To justify this claim let us consider $x+N \in A/N$ with $x+N \neq N$. Also, let us consider the mapping $\pi : A \to A/N$ defined by, $\pi(a) := a+N$. Then if $x+N$ is singular in A/N, then $\langle x+N\rangle$ would be a proper ideal in A/N and so $\pi^{-1}(\langle x+N\rangle)$ would be a proper ideal in A that contains N as a proper subset. However this contradicts the maximality of N. Therefore, $x+N$ must be invertible in A/N. Thus, from Theorem 13.5, we know that A/N is isomorphic to $\mathbb{C}$. Let $\sigma : A/N \to \mathbb{C}$ be an isomorphism that realises this. Then $(\sigma\circ\pi) : A \to \mathbb{C}$ is a multiplicative linear functional (i.e., a homomorphism) and $\mathrm{Ker}(\sigma\circ\pi) = N$. $\square$

By combining the previous three results we get the following useful fact.

Corollary 13.17. *Let $(A, \|\cdot\|)$ be a commutative unital Banach algebra and let $x \in A$. Then x is singular if, and only if, there exists a nonzero multiplicative linear functional $x^* \in A^*$ such that $x \in Ker(x^*)$.*

Exercise 13.18. *Let $(A, \|\cdot\|)$ be a commutative unital Banach algebra and let $x \in A$. Then $\lambda \in \sigma_A(x)$ if, and only if, there exists a nonzero multiplicative linear functional $x^* \in A^*$ such that $\lambda = x^*(x)$.*

Let $(A, \|\cdot\|)$ be a commutative unital Banach algebra and let Δ_A denote the set of all nonzero multiplicative linear functionals on A. The **Gelfand transform** of an element $a \in A$ is the function $\widehat{a} : \Delta_A :\to \mathbb{C}$ defined by, $\widehat{a}(x^*) := x^*(a)$. We know from our work on Banach spaces that $\widehat{a} \in C_{\mathbb{C}}(\Delta_A, \text{weak}^*)$.

Theorem 13.19 (Gelfand, 1941)**.** *If $(A, \|\cdot\|)$ is a commutative unital Banach algebra, then: (i) the mapping $a \mapsto \widehat{a}$ is a unital algebra homomorphism from A into $C_{\mathbb{C}}(\Delta_A)$; (ii) $\sigma_A(a) = \text{range}(\widehat{a}) = \sigma_{C_{\mathbb{C}}(\Delta_A)}(\widehat{a})$ and so $r_A(a) = \|\widehat{a}\|_\infty$ and (iii) $\widehat{A}$ is a subalgebra of $C_{\mathbb{C}}(\Delta_A)$ that contains all the constant functions and separates the points of Δ_A.*

Proof. Consider the mapping $a \mapsto \widehat{a}$ from A into $C_{\mathbb{C}}(\Delta_A)$. As mentioned above we know that this mapping is well-defined, i.e., $\widehat{a} \in C_{\mathbb{C}}(\Delta_A, \text{weak}^*)$ for all $a \in A$.

(i) Now, $\widehat{\mathbf{1}_A}(x^*) = x^*(\mathbf{1}_A) = 1$ for all $x^* \in \Delta_A$ since $\Delta_A \subseteq S(A)$. Therefore, $\widehat{\mathbf{1}_A} = \mathbf{1}_{C(\Delta_A)}$. Next, suppose that $x, y \in A$ and $\lambda \in \mathbb{C}$, then for each $x^* \in \Delta_A$,

$$\widehat{x+y}(x^*) = x^*(x+y) = x^*(x) + x^*(y) = \widehat{x}(x^*) + \widehat{y}(x^*),$$

$$\widehat{\lambda x}(x^*) = x^*(\lambda x) = \lambda x^*(x) = \lambda \widehat{x}(x^*)$$

and

$$\widehat{xy}(x^*) = x^*(xy) = x^*(x)x^*(y) = \widehat{x}(x^*)\widehat{y}(x^*).$$

Therefore, $\widehat{x+y} = \widehat{x} + \widehat{y}$, $\widehat{\lambda x} = \lambda\widehat{x}$ and $\widehat{xy} = \widehat{x}\widehat{y}$. This shows that $a \mapsto \widehat{a}$ is a unital algebra homomorphism.

(ii) This follows from the Exercise 13.18.

The first part of (iii) follows from the fact that $a \mapsto \widehat{a}$ is a unital algebra homomorphism. To show that $\widehat{A}$ separates the points of Δ_A we simply note that if $x^*, y^* \in \Delta_A$ and $x^* \neq y^*$ then there exists an $a \in A$ such that $x^*(a) \neq y^*(a)$. Therefore,

$$\widehat{a}(x^*) = x^*(a) \neq y^*(a) = \widehat{a}(y^*).$$

This completes the proof. $\square$

Application

Let $(A, \|\cdot\|)$ be the commutative unital Banach algebra $\ell^1(\mathbb{Z})$ under convolution. For each $z \in \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$, there is a nonzero homomorphism $f_z : A \to \mathbb{C}$ such that

$$f_z(a) := \sum_{n\in\mathbb{Z}} a(n)z^n \quad \text{for all } a \in \ell^1(\mathbb{Z}).$$

That this defines a homomorphism is not obvious and relies upon careful handling of absolutely convergent series. In fact every $g \in \Delta_A$ has the form f_z for some $z \in \mathbb{T}$. To see this, for each $n \in \mathbb{Z}$, define $e_n \in \ell^1(\mathbb{Z})$ by

$$e_n(k) := \begin{cases} 0 & \text{for } k \neq n; \\ 1 & \text{for } k = n. \end{cases}$$

Observe that e_1 and its inverse e_{-1} generate A, in the sense that A is the smallest Banach algebra that contains e_1 and e_{-1}. Therefore, if $g, h \in \Delta_A$ and $g(e_1) = h(e_1)$, then $g = h$ since if $g(e_1) = h(e_1)$, then

$$g(e_{-1}) = g(e_1^{-1}) = g(e_1)^{-1} = h(e_1)^{-1} = h(e_1^{-1}) = h(e_{-1})$$

and $\{a \in A : g(a) = h(a)\}$ is a Banach subalgebra of A.

Now note that (i) $g(e_n) \in \mathbb{T}$ for all $n \in \mathbb{Z}$ and all $g \in \Delta_A$ and (ii) $f_z(e_1) = z$ for all $z \in \mathbb{T}$. Therefore $f_{g(e_1)}(e_1) = g(e_1)$ for every $g \in \Delta_A$ and so $f_{g(e_1)} = g$ for every $g \in \Delta_A$. Thus, $z \mapsto f_z$ is a bijection from $\mathbb{T}$ onto Δ_A, with inverse given by, $g \mapsto g(e_1)$. Since, (i) $g \mapsto g(e_1)$ is continuous, by the definition of the weak* topology on Δ_A, (ii) $g \mapsto g(e_1)$ is a bijection from Δ_A onto $\mathbb{T}$, (iii) $\mathbb{T}$ is Hausdorff and (iv) Δ_A is compact, it follows that $g \mapsto g(e_1)$ is a homeomorphism. Therefore, $\pi : \mathbb{T} \to \Delta_A$ defined by, $\pi(z) := f_z$ is a homeomorphism. [Since π is the inverse of $g \mapsto g(e_1)$]. Hence, $\pi^* : C(\Delta_A) \to C(\mathbb{T})$ defined by,

$$\pi^*(g)(z) := (g \circ \pi)(z) = g(f_z) \text{ for all } z \in \mathbb{T},$$

is an Banach algebra isomorphism. In particular, if $a := (a(n) : n \in \mathbb{Z}) \in \ell^1(\mathbb{Z})$, then $\widehat{a} \in C(\Delta_A)$ and

$$\pi^*(\widehat{a})(z) = \widehat{a}(f_z) = f_z(a) = \sum_{n \in \mathbb{Z}} a(n) z^n.$$

If $f \in C(\mathbb{T})$ and $f = \pi^*(\widehat{a})$ for some $a \in \ell^1(\mathbb{Z})$, then we can recover $a(n)$ as the n^{th} Fourier coefficient of f. This is,

$$a(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta} \, \mathrm{d}\theta \quad \text{for each } n \in \mathbb{Z}.$$

The algebra $A = \ell^1(\mathbb{Z})$ is often called the **algebra of absolutely convergent Fourier series** because a continuous function $f \in C(\mathbb{T})$ has the form $\pi^*(\widehat{a})$ for some $a \in \ell^1(\mathbb{Z})$ if, and only if, the Fourier coefficients of f form an ℓ^1 sequence on $\mathbb{Z}$. This relies upon the fact that if two continuous functions on $\mathbb{T}$ possess the same Fourier coefficients, then they are equal.

Theorem 13.20 (Wiener)**.** *If f is a unit in $(C(\mathbb{T}), \|\cdot\|_\infty)$, i.e., $0 \notin f(\mathbb{T})$ and has an absolutely convergent Fourier series, then so does $1/f$.*

Proof. (Gelfand) Let $a(n)$ denote the n^{th} Fourier coefficient of f so that $a \in \ell^1(\mathbb{Z})$ by hypothesis. Then $\pi^*(\widehat{a}) \in C(\mathbb{T})$ has the same Fourier coefficients as f, hence equals f. Thus f non-vanishing says that $\pi^*(\widehat{a})$ is a unit in $C(\mathbb{T})$ which in turn implies that

$\widehat{a} \in C(\Delta_{\ell^1(\mathbb{Z})})$ is a unit in $C(\Delta_{\ell^1(\mathbb{Z})})$ and so, by the Gelfand Theorem, a is a unit in $\ell^1(\mathbb{Z})$. But then, $\pi^*(\widehat{a^{-1}})$ is an inverse of $\pi^*(\widehat{a}) = f$, and so

$$(1/f)(z) = \pi^*(\widehat{a^{-1}})(z) = \sum_{n\in\mathbb{Z}} a^{-1}(n) z^n.$$

This shows that $1/f$ has an absolutely convergent Fourier series. $\square$

Chapter 14

C^*-algebras

Given an algebra A over $\mathbb{C}$, an operation $x \mapsto x^*$ on A which satisfies the properties:

(i) $(x+y)^* = x^* + y^*$ for all $x, y \in A$;

(ii) $(\lambda x)^* = \overline{\lambda} x^*$ for all $\lambda \in \mathbb{C}$ and $x \in A$;

(iii) $(xy)^* = y^* x^*$ for all $x, y \in A$;

(iv) $x^{**} = x$ for all $x \in A$.

is called an **involution** on A. An algebra A with an involution $*$ is called a **$*$-algebra**. A Banach algebra $(A, \|\cdot\|)$ with an involution $*$ that is related to the norm by the equation

(v) $\|xx^*\| = \|x\|^2$ for all $x \in A$

is called a **C^*-algebra**. This last requirement of the norm is called the **C^*-condition**.

Exercise 14.1. *Show that in a C^*-algebra $(A, \|\cdot\|)$, $\|x\| = \|x^*\|$ for all $x \in A$.*
Hint: $\|x\|^2 = \|xx^*\| \leqslant \|x\|\|x^*\|$.

Example 14.2. *(a) Let $(H, \langle\cdot,\cdot\rangle)$ be a nontrivial Hilbert space. Then $B(H)$ is a C^*-algebra, the involution being the adjoint operation; (b) Let K be a compact Hausdorff space, then $(C_{\mathbb{C}}(K), \|\cdot\|_\infty)$ is a C^*-algebra, the involution being pointwise conjugation.*

We shall say that an element x of a C^*-algebra A is **normal** if $x^*x = xx^*$ i.e., if x commutes with its adjoint. Moreover, we shall say that an element $x \in A$ is **self-adjoint** if $x = x^*$. Clearly every self-adjoint element is normal.

We shall let A_{sa} denote the set of self-adjoint elements of A. Note that if $(A, \|\cdot\|)$ is a unital C^*-algebra and $a \in A$, then $a\mathbf{1}_A^* = (\mathbf{1}_A a^*)^* = (a^*)^* = a$ and similarly $\mathbf{1}_A^* a = a$. By the uniqueness of the multiplicative identity, it follows that $\mathbf{1}_A = \mathbf{1}_A^*$ and so $\mathbf{1}_A$ is self-adjoint.

Exercise 14.3. *Let $(A, \|\cdot\|)$ be a unital C^*-algebra. Show that:*

(i) $0^* = 0$;

(ii) $x \in A$ is a unit if, and only if, x^* is a unit;

(iii) If $x \in A$ is a unit, then $(x^*)^{-1} = (x^{-1})^*$;

(iv) If $x \in A$, then $\sigma_A(x^*) = \{\overline{\lambda} : \lambda \in \sigma_A(x)\}$.

Lemma 14.4. *If $(A, \|\cdot\|)$ is a C^*-algebra and $a \in A$, then there exist unique self-adjoint elements $b, c \in A$ such that (i) $a = b + ic$ and (ii) $\|b\|, \|c\| \leqslant \|a\|$.*

Proof. Note that $\frac{1}{2}(a+a^*)$ and $\frac{-i}{2}(a-a^*)$ are self-adjoint and $a = \frac{1}{2}(a+a^*) + i\frac{-i}{2}(a-a^*)$. This shows existence. Suppose $a = b + ic$ where $b, c \in A_{sa}$ then $a^* = b - ic$. From these equations we get $b = \frac{1}{2}(a + a^*)$ and $c = \frac{-i}{2}(a - a^*)$. This shows uniqueness.

Using the triangle inequality, $\|b\| = \|\frac{1}{2}(a + a^*)\| \leqslant \frac{1}{2}(\|a\| + \|a^*\|) = \|a\|$ and similarly $\|c\| \leqslant \|a\|$. $\square$

Lemma 14.5. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a state on A. Then $f(a^*) = \overline{f(a)}$ for all $a \in A$. In particular, if f is a multiplicative linear functional on A, then $f(a^*) = \overline{f(a)}$ for all $a \in A$*

Proof. Let $a \in A$ be self-adjoint. Then $f(a) = \alpha + i\beta$ for $\alpha, \beta \in \mathbb{R}$. For each $\lambda \in \mathbb{R}$ consider $b_\lambda := a + i\lambda \mathbf{1}_A$. Note that $\|f\| = 1$ so,

$$|f(b_\lambda)|^2 \leqslant \|b_\lambda\|^2 = \|b_\lambda^* b_\lambda\| = \|(a - i\lambda\mathbf{1}_A)(a + i\lambda\mathbf{1}_A)\| \leqslant \|a\|^2 + \lambda^2.$$

On the other hand from the definition of b_λ,

$$|f(b_\lambda)|^2 = |f(a) + i\lambda f(\mathbf{1}_A)|^2 = |\alpha + i(\beta + \lambda)|^2 = \alpha^2 + \beta^2 + \lambda^2 + 2\lambda\beta.$$

Putting this together gives $\alpha^2 + \beta^2 + 2\beta\lambda \leqslant \|a\|^2$ for all $\lambda \in \mathbb{R}$. But this is impossible unless $\beta = 0$. Thus $f(a) \in \mathbb{R}$. Now in general if $a \in A$ then $a = b + ic$ for $b, c \in A_{sa}$. So,

$$\overline{f(a)} = \overline{f(b + ic)} = \overline{f(b) + if(c)} = f(b) - if(c) = f(b - ic) = f(a^*). \quad \square$$

Corollary 14.6. *Let $(A, \|\cdot\|)$ be a unital C^*-algebra and let f be a state, and in particular a nonzero multiplicative linear functional on A. If $a \in A$ is self-adjoint, then $f(a) \in \mathbb{R}$.*

Exercise 14.7. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and $a \in A$ is normal. Show that a^{2^n} is normal for all $n \in \mathbb{N}$.*

Lemma 14.8. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and $a \in A$ is normal. Then $r_A(a) = \|a\|$.*

Proof. Let a be a normal element of A. Note that $(a^2)^* = a^*a^* = (a^*)^2$. Then,

$$\|a^2\|^2 = \|a^2(a^2)^*\| = \|a^2(a^*)^2\| = \|(aa^*)(aa^*)\| = \|aa^*\|^2 = \|a\|^4.$$

Now proceeding inductively and noting that a^{2^n} is normal for all $n \in \mathbb{N}$ we see that $\|a^{2^k}\| = \|a\|^{2^k}$ for all $k \in \mathbb{N}$. Hence using the spectral radius formula,

$$r_A(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}} = \lim_{k\to\infty} \|a^{2^k}\|^{\frac{1}{2^k}} = \|a\|. \quad \square$$

Theorem 14.9. *If $(A, \|\cdot\|)$ is a unital C^*-algebra, then $\|a\| = \sqrt{r_A(aa^*)}$. In particular, the norm on A is completely determined by the algebraic structure on A.*

Proof. Let $(A, \|\cdot\|)$ be a unital C^*-algebra and let $a \in A$. Then $r_A(aa^*) = \|aa^*\| = \|a\|^2$, since aa^* is self-adjoint and hence normal. Therefore, $\|a\| = \sqrt{r_A(aa^*)}$. Now the right-hand side of this equation is solely determined by the algebraic structure of A. $\square$

The next corollary shows that unital $*$-homomorphisms between unital C^*-algebras are automatically bounded and hence continuous.

Corollary 14.10. *Suppose that $(A, \|\cdot\|)$ and $(B, \|\!|\cdot|\!\|)$ are unital C^*-algebras and $\pi : A \to B$ is a unital $*$-homomorphism. Then $\|\!|\pi(a)|\!\| \leqslant \|a\|$ for all $a \in A$.*

Proof. For $a \in A$ we have $\sigma_B(\pi(a^*a)) \subseteq \sigma_A(a^*a)$ and so

$$\|\!|\pi(a)|\!\| = \sqrt{r_B(\pi(a)^*\pi(a))} = \sqrt{r_B(\pi(a^*a))} \leqslant \sqrt{r_A(a^*a)} = \|a\|. \quad \square$$

Theorem 14.11 (Commutative Gelfand-Naimark)**.** *Suppose $(A, \|\cdot\|)$ is a nonzero commutative unital C^*-algebra. Then the Gelfand transform $a \mapsto \widehat{a}$ is an isometric $*$-isomorphism from A onto $C(\Delta_A)$.*

Proof. We know $a \to \widehat{a}$ preserves scalar multiplication, addition and multiplication. Further for $a \in A$ and $f \in \Delta_A$ using Lemma 14.5,

$$\widehat{a^*}(f) = f(a^*) = \overline{f(a)} = \overline{\widehat{a}(f)}.$$

It follows the Gelfand transform is a $*$-homomorphism. As A is commutative every element of A is normal. Hence,

$$\|a\| = r_A(a) = \|\widehat{a}\|_\infty$$

for all $a \in A$. It follows the Gelfand transform is isometric, hence injective and $\widehat{A}$ is closed. Finally as $\widehat{A}$ is a closed self-adjoint subalgebra of $C(\Delta_A)$ that contains all the constant function, it follows from the Stone-Weierstrass Theorem that $\widehat{A} = C(\Delta_A)$ and so the Gelfand transform is surjective. This completes the proof. $\square$

Even though the Commutative Gelfand-Naimark Theorem only applies to commutative unital C^*-algebras we shall later see it can be useful even if the C^*-algebra is not commutative or have a multiplicative identity. Note that if $(A, \|\cdot\|)$ is a C^*-algebra, S is a set and for each $s \in S$, B_s is a C^*-subalgebra of A, then the intersection $B := \bigcap_{s \in S} B_s$ is also a C^*-subalgebra of A. If $S \subseteq A$ we shall let $C(S)$ denote the smallest C^*-subalgebra of A containing S. That is, $C(S)$ is the intersection of all C^*-algebras containing S. If $\{a_1, a_2, ..., a_n\} \subseteq A$, then we will write $C(a_1, a_2, ..., a_n)$ instead of $C(\{a_1, a_2, ..., a_n\})$.

Lemma 14.12. *Suppose that $(A, \|\cdot\|)$ is a C^*-algebra and $a \in A$. Then*

$$Comm(a) := \{b \in A : ab = ba \text{ and } ab^* = b^*a\}$$

is a C^*-subalgebra of $(A, \|\cdot\|)$.

Proof. It is easy to see that $\mathrm{Comm}(a)$ is a subspace of A that is closed under multiplication and the involution. Further, suppose $(b_n : n \in \mathbb{N})$ is a sequence in $\mathrm{Comm}(a)$ converging to $b \in A$. Then $ab_n = b_n a$ for all $n \in \mathbb{N}$. Therefore, as multiplication is continuous, $ab = \lim_{n\to\infty} ab_n = \lim_{n\to\infty} b_n a = ba$. As the involution is continuous $ab^* = b^*a$ also. It follows $b \in \mathrm{Comm}(a)$. Hence, $\mathrm{Comm}(a)$ is closed in the norm topology and so is a C^*-subalgebra of A. This completes the proof. □

Lemma 14.13. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and $a \in A$ is normal. Then $C(a, \mathbf{1}_A)$ is a commutative C^*-algebra.*

Proof. Consider $\mathrm{Comm}(a)$. As a is normal $aa^* = a^*a$ and so $\mathrm{Comm}(a)$ is a C^*-algebra containing $\mathbf{1}_A$ and a. For $b \in \mathrm{Comm}(a)$, $ab = ba$ and $ab^* = b^*a$. That is, $ba = ab$ and $ba^* = a^*b$ and so $\mathbf{1}_A, a \in \mathrm{Comm}(b)$. Define $C := \bigcap_{b\in\mathrm{Comm}(a)} \mathrm{Comm}(b)$. Then C is a C^* algebra containing $\mathbf{1}_A$ and a. It follows that $C(a, \mathbf{1}_A) \subseteq C$. Further, since $a \in \mathrm{Comm}(a)$, $C(a, \mathbf{1}_A) \subseteq C \subseteq \mathrm{Comm}(a)$. Now, if $c, d \in C(a, \mathbf{1}_A)$, then $c \in \mathrm{Comm}(a)$ and $d \in C(a, \mathbf{1}_A) \subseteq C \subseteq \mathrm{Comm}(c)$. Therefore, it follows that $cd = dc$, and so $C(a, \mathbf{1}_A)$ is commutative. □

The Commutative Gelfand-Naimark Theorem allows us to construct a continuous **functional calculus**. If $(A, \|\cdot\|)$ is a unital C^*-algebra and $a \in A$ is normal, then $C(a, \mathbf{1}_A)$ is a commutative unital C^*-algebra. Let f be a function continuous on $\sigma_A(a)$. Then $f \circ \widehat{a} \in C(\Delta_A)$. We let $f(a)$ denote the unique element of $C(a, \mathbf{1}_A)$ such that $\widehat{f(a)} = f \circ \widehat{a}$. This construction has many desirable properties.

Corollary 14.14. *Suppose that $(A, \|\cdot\|)$ is unital C^*-algebra and $a \in A$ is self-adjoint. Then $\sigma_A(a) \subseteq \mathbb{R}$*

Proof. Consider the commutative unital C^*-algebra $C(a, \mathbf{1}_A)$. As $a = a^*$, applying the Commutative Gelfand-Naimark Theorem we get $\widehat{a} = \widehat{a^*} = \overline{\widehat{a}}$ and so $\mathrm{range}(\widehat{a}) \subseteq \mathbb{R}$. Hence, $\sigma_A(a) \subseteq \sigma_{C(a,\mathbf{1}_A)}(a) = \mathrm{range}(\widehat{a}) \subseteq \mathbb{R}$. □

Lemma 14.15. *Suppose that $(A, \|\cdot\|)$ is a unital C^*-algebra with identity $\mathbf{1}_A$ and B is a C^*-subalgebra of A with $\mathbf{1}_A \in B$. Then for $a \in B$, a is a unit in B if, and only if, a is a unit in A. In particular $\sigma_B(a) = \sigma_A(a)$.*

Proof. Suppose first that a is self-adjoint. Then $\sigma_B(a) \subseteq \mathbb{R}$ by Corollary 14.14. Then $\sigma_B(a)$ a closed subset of $\mathbb{C}$ with empty interior so $\partial\sigma_B(a) = \sigma_B(a)$. Then, by Proposition 13.2

$$\sigma_B(a) = \partial\sigma_B(a) \subseteq \sigma_A(a) \subseteq \sigma_B(a)$$

and so $\sigma_B(a) = \sigma_A(a)$. Noting $0 \in \sigma_A(a)$ if, and only if, a is singular, the result follows for self-adjoint elements. Now for arbitrary $a \in B$, suppose a is a unit in A. Then a^*a is also a unit in A. But as a^*a is self-adjoint from the special case previously proved a^*a is a unit in B and so $(a^*a)^{-1} \in B$. Then,

$$a^{-1} = a^{-1}\mathbf{1}_A = a^{-1}(aa^{-1})^* = a^{-1}((a^{-1})^*a^*) = (a^{-1}(a^{-1})^*)a^* = (a^*a)^{-1}a^*$$

so a^{-1} a product of elements in B and so $a^{-1} \in B$. It follows a is invertible in B. The reverse implication is obvious as $B^{-1} \subseteq A^{-1}$. □

Chapter 15

Positive elements

Our first goal in this section is to show that for a self-adjoint element a of a unital C^*-algebra $(A, \|\cdot\|)$, $\sigma_A(a) \subseteq [0, \infty)$ if, and only if, $a = bb^*$ for some $b \in A$.

Lemma 15.1. *Suppose that a is a normal element of a unital C^*-algebra $(A, \|\cdot\|)$. If $\lambda \in \mathbb{C}$ and $r \geqslant 0$, then $\sigma_A(a) \subseteq B[\lambda, r]$ if, and only if, $\|a - \lambda \mathbf{1}_A\| \leqslant r$.*

Proof. Suppose that $a \in A$ is normal, $\lambda \in \mathbb{C}$ and $r \geqslant 0$, then

$$\begin{array}{ll} & \sigma_A(a) \subseteq B[\lambda, r] \\ \iff & \sigma_A(a) \subseteq \lambda + B[0, r] \\ \iff & \sigma_A(a) - \lambda \subseteq B[0, r] \\ \iff & \sigma_A(a - \lambda \mathbf{1}_A) \subseteq B[0, r] \\ \iff & r_A(a - \lambda \mathbf{1}_A) \leqslant r \\ \iff & \|a - \lambda \mathbf{1}_A\| \leqslant r \quad \text{since } a - \lambda \mathbf{1} \text{ is normal.} \end{array}$$

This completes the proof. $\square$

Corollary 15.2. *Suppose $\{a_1, a_2, \ldots, a_n\}$ are normal elements of a unital C^*-algebra $(A, \|\cdot\|)$. If $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subseteq \mathbb{C}$ and $\{r_1, r_2, \ldots, r_n\} \subseteq [0, \infty)$ are such that $\sigma_A(a_k) \subseteq B[\lambda_k, r_k]$ for all $1 \leqslant k \leqslant n$, then*

$$\sigma_A(a_1 + a_2 + \cdots + a_n) \subseteq B[(\lambda_1 + \lambda_2 + \cdots + \lambda_n), (r_1 + r_2 + \cdots + r_n)].$$

Proof. Suppose that $\{a_1, a_2, \ldots, a_n\}$ are normal elements of A, $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subseteq \mathbb{C}$ and $\{r_1, r_2, \ldots, r_n\} \subseteq [0, \infty)$. Then,

$$\begin{array}{rcl} 0 & \leqslant & r_A((a_1 + a_2 + \cdots + a_n) - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\mathbf{1}_A) \\ & \leqslant & \|(a_1 + a_2 + \cdots + a_n) - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\mathbf{1}_A\| \\ & = & \|(a_1 - \lambda_1 \mathbf{1}_A) + (a_2 - \lambda_2 \mathbf{1}_A) + \cdots + (a_n - \lambda_n \mathbf{1}_A)\| \\ & \leqslant & \|a_1 - \lambda_1 \mathbf{1}_A\| + \|a_2 - \lambda_2 \mathbf{1}_A\| + \cdots + \|a_n - \lambda_n \mathbf{1}_A\| \\ & \leqslant & r_1 + r_2 + \cdots + r_n, \quad \text{by the Lemma 15.1.} \end{array}$$

Therefore, $\sigma_A((a_1 + a_2 + \cdots + a_n) - (\lambda_1 + \lambda_2 + \cdots \lambda_n)\mathbf{1}_A) \subseteq B[0, r_1 + r_2 + \cdots + r_n]$ and so $\sigma_A(a_1 + a_2 + \cdots + a_n) - (\lambda_1 + \lambda_2 + \cdots + \lambda_n) \subseteq B[0, r_1 + r_2 + \cdots + r_n]$, i.e., $\sigma_A(a_1 + a_2 + \cdots + a_n) \subseteq B[(\lambda_1 + \lambda_2 + \cdots \lambda_n), (r_1 + r_2 + \cdots + r_n)]$. $\square$

Theorem 15.3. *Suppose that a and b are self-adjoint elements of a unital C^*-algebra $(A, \|\cdot\|)$. If $\sigma_A(a) \subseteq [0, \infty)$ and $\sigma_A(b) \subseteq [0, \infty)$, then $\sigma_A(a + b) \subseteq [0, \infty)$.*

Proof. Firstly, note that $a + b$ is self-adjoint and so $\sigma_A(a + b) \subseteq \mathbb{R}$. Let $\lambda_1 := r_1 := \|a\|/2$ and $\lambda_2 := r_2 := \|b\|/2$. Then $\sigma_A(a) \subseteq [0, \|a\|] \subseteq B[\lambda_1, r_1]$ and $\sigma_A(b) \subseteq [0, \|b\|] \subseteq B[\lambda_2, r_2]$. Therefore, by Corollary 15.2, $\sigma_A(a + b) \subseteq B[(\lambda_1 + \lambda_2), (r_1 + r_2)] \cap \mathbb{R} = [0, \|a\| + \|b\|]$. $\square$

Unfortunately, we are still unable to prove the desired result that for any element a of a unital C^*-algebra $(A\|\cdot\|)$, $\sigma_A(a) \subseteq [0, \infty)$ whenever $a = bb^*$ for some $b \in A$. However, we can easily prove the following partial result.

Proposition 15.4. *Suppose that a is any element of a unital C^* algebra $(A, \|\cdot\|)$. Then $\sigma_A(aa^* + a^*a) \subseteq [0, \infty)$.*

Proof. Write a as: $a = x + iy$, where x and y are self-adjoint elements of A. Then,

$$aa^* + a^*a = (x + iy)(x - iy) + (x - iy)(x + iy) = 2(x^2 + y^2).$$

Because x and y are self-adjoint we have, via the Gelfand-Naimark Theorem, applied to $C(x, \mathbf{1}_A)$ and $C(y, \mathbf{1}_A)$, that

$$\sigma_A(x^2) = \sigma_{C(x,\mathbf{1}_A)}(x^2) = \text{range}[(\widehat{x})^2] \subseteq [0, \infty)$$

and

$$\sigma_A(y^2) = \sigma_{C(y,\mathbf{1}_A)}(y^2) = \text{range}[(\widehat{y})^2] \subseteq [0, \infty).$$

Hence, by Theorem 15.3,

$$\sigma_A(aa^* + a^*a) = \sigma_A(2(x^2 + y^2)) = 2\sigma_A(x^2 + y^2) \subseteq [0, \infty).$$

This completes the proof. $\square$

Theorem 15.5 (Square Root Theorem)**.** *Let a be a self-adjoint element of a unital C^*-algebra $(A, \|\cdot\|)$. Then $\sigma_A(a) \subseteq [0, \infty)$ if, and only if, $a = bb^*$ for some $b \in A$.*

Proof. It follows from the Gelfand-Naimark Theorem applied to $C(a, \mathbf{1}_A)$ that if a is a self-adjoint element and $\sigma_A(a) = \sigma_{C(a,\mathbf{1}_A)}(a) \subseteq [0, \infty)$, then there exists a self-adjoint element $b \in C(a, \mathbf{1}_A)$ such that $a = b^2 = bb^*$. This is essentially an application of functional calculus. So we concern ourselves with the converse.

Suppose that $a = bb^*$ for some $b \in A$. In order to obtain a contradiction let us suppose that $\sigma_A(a) = \sigma_{C(a,\mathbf{1}_A)}(a) \not\subseteq [0, \infty)$. We shall first show that this implies that there exists a nonzero element $d \in C(a, \mathbf{1}_A)$ such that $\sigma_A(d^*d) = \sigma_{C(a,\mathbf{1}_A)}(d^*d) \subseteq (-\infty, 0]$. From the Gelfand-Naimark Theorem applied to $C(a, \mathbf{1}_A)$ we have range$(\widehat{a}) = \sigma_{C(a,\mathbf{1}_A)}(a) \not\subseteq [0, \infty)$.

Therefore there exists a "bump" function $g \in C_{\mathbb{C}}(\Delta_{C(a,\mathbf{1}_A)})$ such that: (i) $g : \Delta_{C(a,\mathbf{1}_A)} \to [0,1]$; (ii) $\|g\widehat{a}\|_\infty \neq 0$ and (iii) range$(g\widehat{a}) \subseteq (-\infty, 0]$. For example, $g := \frac{-1}{\|a\|} \min\{\hat{a}, 0\}$. Let $h \in C_{\mathbb{C}}(\Delta_{C(a,\mathbf{1}_A)})$ be defined by, $h(x^*) = \sqrt{g(x^*)}$ for all $x^* \in \Delta_{C(a,\mathbf{1}_A)}$. Next, select $c \in C(a, \mathbf{1}_A)$ so that $\widehat{c} = h$ and note that $c = c^*$ since $h = \overline{h}$. Then,

$$\widehat{cac} = \widehat{c}\widehat{a}\widehat{c} = h\widehat{a}h = h^2\widehat{a} = g\widehat{a}.$$

Therefore, $cac \neq 0$ since $g\widehat{a} \neq 0$ and the Gelfand transform is 1-to-1. Furthermore,

$$\sigma_A(cac) = \sigma_{C(a,\mathbf{1}_A)}(cac) = \text{range}(\widehat{cac}) = \text{range}(g\widehat{a}) \subseteq (-\infty, 0].$$

Let $d := b^*c$, then

$$d^*d = (b^*c)^*(b^*c) = (cb)(b^*c) = c(bb^*)c = cac.$$

Thus, $\sigma_A(d^*d) \subseteq (-\infty, 0]$ and $d \neq 0$, since $cac \neq 0$.

We will now use this d to obtain a contraction. Since $\sigma_A(dd^*) \setminus \{0\} = \sigma_A(d^*d) \setminus \{0\}$ we also have that $\sigma_A(dd^*) \subseteq (-\infty, 0]$. Thus, from Theorem 15.3, $\sigma_A(dd^* + d^*d) \subseteq (-\infty, 0]$. On the other hand, by Proposition 15.4, $\sigma_A(dd^* + d^*d) \subseteq [0, \infty)$, i.e., $\sigma_A(dd^* + d^*d) = \{0\}$. Since $dd^* + d^*d$ is self-adjoint, $\|dd^* + d^*d\| = r_A(dd^* + d^*d) = 0$, i.e., $dd^* = -d^*d$. In particular, this implies that

$$\sigma_A(dd^*) \setminus \{0\} = \sigma_A(-d^*d) \setminus \{0\} = -(\sigma_A(d^*d) \setminus \{0\}) \subseteq [0, \infty)$$

i.e., $\sigma_A(dd^*) \subseteq (-\infty, 0] \cap [0, \infty) = \{0\}$. Thus, $\|d\|^2 = \|dd^*\| = r_A(dd^*) = 0$. However, this contradicts our assumption that $d \neq 0$. Hence, $\sigma_A(a) \subseteq [0, \infty)$. $\square$

Let $(A, \|\cdot\|)$ be a unital C^*-algebra. An element $a \in A$ is said to be **positive** if it is self-adjoint and $\sigma_A(a) \subseteq [0, \infty)$. Or equivalently, by the Square Root Theorem, if $a = bb^*$ for some $b \in A$. We shall denote by A_+ the set of all positive element of A.

If V is a vector space over $\mathbb{R}$ and C is a subset of V such that $C \cap (-C) = \{0\}$ and $\alpha a + \beta b \in C$ for all $x, y \in C$ and $\alpha, \beta \in [0, \infty)$, then we say C is a **cone** of V.

Lemma 15.6. *Suppose V is a vector space over $\mathbb{R}$ and C is a cone of V. If we define a relation on V by, $x \geqslant y$, if $x - y \in C$, then $\geqslant$ is a partial order on V.*

Proof. Note $x - x = 0 \in C$ so $x \geqslant x$. If $x \geqslant y$ and $y \geqslant x$, then $x - y, -(x - y) \in C$ so $x = y$. If $x \geqslant y$ and $y \geqslant z$, then $x - y, y - z \in C$ so $x - z = (x - y) + (y - z) \in C$ so $x \geqslant z$. It follows $\geqslant$ is a partial ordering of V. $\square$

If $(A, \|\cdot\|)$ is a unital C^*-algebra we can regard A_{sa} as a vector space over $\mathbb{R}$ in a natural way. From Theorem 15.3 it is obvious that $\alpha a + \beta b \in A_+$ for all $a, b \in A_+$ and $\alpha, \beta \in [0, \infty)$. Moreover, if $a \in A_+ \cap (-A_+)$, then $\sigma_A(a) = \{0\}$ so as a is self-adjoint and hence normal $\|a\| = r_A(a) = 0$ so $a = 0$. It follows A_+ is a cone of A_{sa} and the relation "$\geqslant$" defined by $a \geqslant b$, if $a - b \in A_+$ is a partial ordering of A_{sa}.

Lemma 15.7. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra. If $a \in A$ is self adjoint, then $-\|a\|\mathbf{1}_A \leqslant a \leqslant \|a\|\mathbf{1}_A$.*

Proof. Suppose $a \in A$ is self-adjoint. Consider $C(a, \mathbf{1}_A)$,

$$\sigma_A(a + \|a\|\mathbf{1}_A) = \sigma_{C(a,\mathbf{1}_A)}(a + \|a\|\mathbf{1}_A) = \mathrm{range}(\widehat{a} + \|a\|\widehat{\mathbf{1}_A}) = \mathrm{range}(\widehat{a}) + \|\widehat{a}\|_\infty \subseteq [0, \infty).$$

So $a + \|a\|\mathbf{1}_A$ is positive. Therefore, $-\|a\|\mathbf{1}_A \leqslant a$. Similarly, $a \leqslant \|a\|\mathbf{1}_A$. $\square$

Sesquilinear Forms

Suppose V is a vector space over $\mathbb{C}$ and $[\cdot,\cdot] : V \times V \to \mathbb{C}$ is a map that is linear in the first variable and conjugate linear in the second variable. That is,

(i) $[w + x, y + z] = [w, y] + [w, z] + [x, y] + [x, z]$ for all $w, x, y, z \in V$,

(ii) $[\alpha x, \beta y] = \alpha\overline{\beta}[x, y]$ for all $x, y \in V$ and $\alpha, \beta \in \mathbb{C}$.

Then we say that $[\cdot,\cdot]$ is a **sesquilinear form**. Further,

(i) if $[x, x] \geqslant 0$ for all $x \in V$, then we say that $[\cdot,\cdot]$ is **positive sesquilinear form**,

(ii) if $[x, y] = \overline{[y, x]}$ for all $x, y \in V$, then we say that $[\cdot,\cdot]$ is a **hermitian sesquilinear form**,

(iii) if $[\cdot,\cdot]$ is positive and $[x, x] = 0 \implies x = 0$, then we say that $[\cdot,\cdot]$ is a **positive definite sesquilinear form**.

Note that if V is a vector space and $[\cdot,\cdot]$ is a sesquilinear form on V, then for any $x \in V$,

$$2[x, 0] = [2x, 0] = [x + x, 0 + 0] = [x, 0] + [x, 0] + [x, 0] + [x, 0] = 4[x, 0]$$

and so $[x, 0] = 0$. Similarly, it follows that $[0, x] = 0$.

Let $(A, \|\cdot\|)$ be a unital C^*-algebra. Suppose f is a linear functional on A. We say f is a **positive linear functional** if $f(a) \geqslant 0$ for all $a \in A_+$. Note that positive linear functionals respect the ordering on A_{sa}. If $a \geqslant b$, then $a - b$ is positive and so $f(a - b) \geqslant 0$. Therefore, $f(a) - f(b) \geqslant 0$ and so $f(a) \geqslant f(b)$.

Note that if K is a compact Hausdorff topological space and μ is a positive Borel measure on K, then $x^* : C_{\mathbb{C}}(K) \to \mathbb{C}$ defined by, $x^*(f) := \int_K f \,\mathrm{d}\mu$ for all $f \in C_{\mathbb{C}}(K)$, is a positive functional on $C_{\mathbb{C}}(K)$. Furthermore, if $Tr : M_n(\mathbb{C}) \to \mathbb{C}$ is defined by $Tr((a_{ij})) := \sum_{i=1}^n a_{ii}$, for all $(a_{ij}) \in M_n(\mathbb{C})$, then Tr is a positive functional on $M_n(\mathbb{C})$.

Lemma 15.8. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A. Then f is bounded.*

Proof. Consider $a \in A_{sa}$. Then $-\|a\|\mathbf{1}_A \leqslant a \leqslant |a\|\mathbf{1}_A$ so $-\|a\|f(\mathbf{1}_A) \leqslant f(a) \leqslant \|a\|f(\mathbf{1}_A)$. Hence, $|f(a)| \leqslant \|a\|f(\mathbf{1}_A)$. In general, if $a \in A$, then there exist self-adjoint elements $b, c \in A$ such that $a = b + ic$, $\|b\| \leqslant \|a\|$ and $\|c\| \leqslant \|a\|$. Then,

$$|f(a)| = |f(b + ic)| \leqslant |f(b)| + |f(c)| \leqslant \|b\|f(\mathbf{1}_A) + \|c\|f(\mathbf{1}_A) \leqslant 2f(\mathbf{1}_A)\|a\|.$$

Which shows f is bounded and in particular $\|f\| \leqslant 2f(\mathbf{1}_A)$. $\square$

Example 15.9. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A. Define $[a,b]_f := f(b^*a)$. Then $[\cdot,\cdot]_f$ is a positive sesquilinear form. In fact $[a,b] := f(ab^*)$ is also a positive sesquilinear form, but we will not use this latter.*

Lemma 15.10. *Suppose V is a vector space over $\mathbb{C}$ and $[\cdot,\cdot] : V \times V \to \mathbb{C}$ is a positive sesquilinear form. Then $[\cdot,\cdot]$ is a hermitian sesquilinear form.*

Proof. Suppose $x, y \in V$ and $\lambda \in \mathbb{C}$ then,

$$0 \leqslant [x + \lambda y, x + \lambda y] = [x,x] + |\lambda|^2[y,y] + \lambda[y,x] + \overline{\lambda}[x,y].$$

As $[x,x] + |\lambda|^2[y,y] \in \mathbb{R}$ it follows that $\mathrm{Im}(\lambda[y,x] + \overline{\lambda}[x,y]) = 0$. Setting $\lambda = 1$ and $\lambda = i$ shows that $\mathrm{Im}[x,y] = -\mathrm{Im}[y,x]$ and $\mathrm{Re}[x,y] = \mathrm{Re}[y,x]$. Thus, $[\cdot,\cdot]$ is hermitian. □

If $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A, then

$$f(a^*) = f(a^*\mathbf{1}_A) = [\mathbf{1}_A, a]_f = \overline{[a, \mathbf{1}_A]_f} = \overline{f(\mathbf{1}_A^* a)} = \overline{f(a)}.$$

This shows positive linear functionals preserve the involution. Recall also that a positive definite sesquilinear form is an inner product.

Lemma 15.11. *Suppose V is a vector space and $[\cdot,\cdot] : V \times V \to \mathbb{C}$ is a positive sesquilinear form. If $y \in V$ is such that $[y,y] = 0$, then $[y,x] = [x,y] = 0$ for all $x \in V$.*

Proof. Recall from previously that $[\cdot,\cdot]$ is hermitian, as $[\cdot,\cdot]$ is positive. Let $x \in V$ and set $\lambda := -t[x,y]$ for $t \in \mathbb{R}$. Then

$$0 \leqslant [x + \lambda y, x + \lambda y] = [x,x] + \overline{\lambda}[x,y] + \lambda[y,x] = [x,x] - 2t\big|[x,y]\big|^2.$$

So if $[x,y] \neq 0$, then $[x,x] - 2t\big|[x,y]\big|^2$ is negative for large enough $t \in \mathbb{R}$. Therefore, it follows that $[x,y] = 0$. As $[\cdot,\cdot]$ is hermitian it also follows that $[y,x] = 0$. □

The next lemma gives a version of the Cauchy-Schwarz inequality for positive sesquilinear forms.

Lemma 15.12 (Cauchy-Schwarz inequality)**.** *Suppose V is a vector space and $[\cdot,\cdot] : V \times V \to \mathbb{C}$ is a positive sesquilinear form. Then $\big|[x,y]\big|^2 \leqslant [x,x][y,y]$ for all $x, y \in V$.*

Proof. Let $x, y \in V$ note that as $[\cdot,\cdot]$ is positive, $[\cdot,\cdot]$ is hermitian. Consider first the case when $[y,y] = 0$. Then, by above, $[x,y] = 0$ also and so $\big|[x,y]\big|^2 \leqslant [x,x][y,y]$ holds. Now suppose that $[y,y] \neq 0$. Set $\alpha := [y,y]$ and $\beta := -[x,y]$. Then,

$$\begin{aligned} 0 \leqslant [\alpha x + \beta y, \alpha x + \beta y] &= |\alpha|^2[x,x] + |\beta|^2[y,y] + \alpha\overline{\beta}[x,y] + \beta\overline{\alpha}[y,x] \\ &= [y,y]^2[x,x] - [y,y]\big|[x,y]\big|^2 \end{aligned}$$

so rearranging and dividing by $[y,y]$ we get that $\big|[x,y]\big|^2 \leqslant [x,x][y,y]$. □

If $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A, then by applying this result to $[a,b]_f$ we get that $|f(b^*a)|^2 \leqslant f(a^*a)f(b^*b)$ for all $a, b \in A$. We call this the **Cauchy-Schwarz inequality for positive linear functionals**.

Theorem 15.13. *Suppose V is a vector space and $[\cdot,\cdot] : V \times V \to \mathbb{C}$ is a positive sesquilinear form. Then: (i) $N := \{x \in V : [x,x] = 0\}$ is a subspace of V; (ii) the map $\langle\cdot,\cdot\rangle : V/N \times V/N \to \mathbb{C}$ defined by $\langle x+N, y+N\rangle := [x,y]$ is a well defined inner product on V/N.*

Proof. First note $0 \in N$. If $x, y \in N$ and $\alpha, \beta \in \mathbb{C}$, then by the previous lemma, $[x,y] = [y,x] = 0$ and so,

$$[\alpha x + \beta y, \alpha x + \beta y] = |\alpha|^2[x,x] + |\beta|^2[y,y] + \alpha\overline{\beta}[x,y] + \beta\overline{\alpha}[y,x] = 0$$

This shows $\alpha x + \beta y \in N$. It follows N is a subspace of V. If $x_1, x_2, y_1, y_2 \in V$ are such that $x_1 + N = x_2 + N$ and $y_1 + N = y_2 + N$, then $x_1 - x_2, y_1 - y_2 \in N$ so, by the previous lemma, $[x_2, y_1 - y_2] = 0$, $[x_1 - x_2, y_2] = 0$ and $[x_1 - x_2, y_1 - y_2] = 0$. Then,

$$\begin{aligned}
[x_1, y_1] &= [x_2 + (x_1 - x_2), y_2 + (y_1 - y_2)] \\
&= [x_2, y_2] + [x_2, y_1 - y_2] + [x_1 - x_2, y_2] + [x_1 - x_2, y_1 - y_2] \\
&= [x_2, y_2].
\end{aligned}$$

This shows $\langle\cdot,\cdot\rangle$ is well defined. The fact that $\langle\cdot,\cdot\rangle$ is linear in the first variable and conjugate linear in the second is easily verified as $[\cdot,\cdot]$ is a positive (hence hermitian) sesquilinear form. As $[\cdot,\cdot]$ is positive it follows $\langle x+N, x+N\rangle := [x,x] \geqslant 0$ for all $x + N \in A/N$. Finally, if $\langle x+N, x+N\rangle = 0$, then $[x,x] = 0$ and so $x \in N$, that is, $x + N = 0$. It follows $\langle\cdot,\cdot\rangle$ is positive definite. □

More on Positive Linear Functionals

Lemma 15.14. *Suppose that f is a state on a unital C^*-algebra $(A, \|\cdot\|)$ and a is a normal element of A. If $\lambda \in \mathbb{C}$, and $r \geqslant 0$ are such that $\sigma_A(a) \subseteq B[\lambda, r]$, then $f(a) \in B[\lambda, r]$.*

Proof. From Lemma 15.1 we know that $\|a - \lambda\mathbf{1}_A\| \leqslant r$. Therefore,

$$|f(a) - \lambda| = |f(a) - f(\lambda\mathbf{1}_A)| = |f(a - \lambda\mathbf{1}_A)| \leqslant \|f\|\|a - \lambda\mathbf{1}_A\| \leqslant r.$$

This completes the proof. □

Theorem 15.15. *Suppose that f is a linear functional on a unital C^*-algebra $(A, \|\cdot\|)$. Then f is a positive functional if, and only if, $\|f\| = f(\mathbf{1}_A)$.*

Proof. Suppose first that f is a positive functional on a unital C^*-algebra $(A, \|\cdot\|)$. If $f = 0$, then the result is obvious, so suppose that $f \neq 0$. By Lemma 15.8 we know that f is bounded. Consider $a \in A$ with $\|a\| = 1$. Then

$$\begin{aligned} |f(a)|^2 &= |f(\mathbf{1}_A^* a)|^2 \quad \text{since } \mathbf{1}_A^* = \mathbf{1}_A \\ &\leqslant f(\mathbf{1}_A^* \mathbf{1}_A) f(a^* a) \quad \text{by Cauchy-Schwarz inequality} \\ &= f(\mathbf{1}_A) f(a^* a) \quad \text{since } \mathbf{1}_A^* = \mathbf{1}_A \\ &\leqslant f(\mathbf{1}_A)\|f\|\|a^* a\| = f(\mathbf{1}_A)\|f\|\|a\|^2 = f(\mathbf{1}_A)\|f\|. \end{aligned}$$

Therefore,

$$\|f\|^2 = \left[\sup_{\|a\|=1} |f(a)|\right]^2 = \sup_{\|a\|=1} |f(a)|^2 \leqslant f(\mathbf{1}_A)\|f\| \leqslant \|f\|^2 \quad \text{since } \|\mathbf{1}_A\| = 1$$

and so $f(\mathbf{1}_A) = \|f\|$, since $\|f\| \neq 0$. Conversely, suppose that $\|f\| = f(\mathbf{1}_A)$. If $\|f\| = 0$, then the result is obvious, so suppose that $\|f\| \neq 0$. Let $g := f/\|f\|$. Then $\|g\| = g(\mathbf{1}_A) = 1$ and so g is a state on A. Let a be any element of A. From the section on C^*-algebras we already know that $g(aa^*) \in \mathbb{R}$ since $(aa^*)^* = aa^*$. We now show that $0 \leqslant g(aa^*)$. Let $\lambda := r := \|a\|^2/2$. Then $\sigma_A(aa^*) \subseteq [0, \|a\|^2] \subseteq B[\lambda, r]$. Therefore, by Lemma 15.14, $g(aa^*) \in B[\lambda, r] \cap \mathbb{R} = [0, \|a\|^2]$. Hence, g is a positive functional. Since $f = \|f\|g$ it follows that f is a positive functional as well. $\square$

Corollary 15.16. *Suppose that f is a positive linear functional on a unital C^*-algebra. If either $\|f\| = 1$ or $f(\mathbf{1}_A) = 1$, then f is a state on A.*

Proof. By Theorem 15.15, $\|f\| = f(\mathbf{1}_A)$ and so the result follows immediately. $\square$

The next lemma establishes a technical inequality that will be used later.

Lemma 15.17. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A. Then for $a, b \in A$, $f((ab)^* ab) \leqslant f(b^* b)\|a\|^2$.*

Proof. If $f(b^* b) = 0$, then $f((ab)^* ab) = f(b^* a^* ab) = f(b^*(a^* ab)) = [a^* ab, b]_f = 0$, by the Cauchy-Schwarz inequality, and so the inequality holds. Suppose $f(b^* b) \neq 0$. Define $g : A \to \mathbb{C}$ by $g(c) := \frac{f(b^* cb)}{f(b^* b)}$ for all $c \in A$. Then g is linear and

$$g(c^* c) = \frac{f(b^* c^* cb)}{f(b^* b)} = \frac{f((cb)^* cb)}{f(b^* b)} \geqslant 0 \quad \text{for all } c \in A.$$

Therefore, g is positive. Moreover, $g(\mathbf{1}_A) = 1$. Hence, g is a state. Therefore,

$$|g(a^* a)| \leqslant \|a^* a\| = \|a\|^2 \quad \text{for all } a \in A$$

and so $f((ab)^* ab) \leqslant f(b^* b)\|a\|^2$ for all $a, b \in A$. $\square$

The GNS Construction

If $(A, \|\cdot\|)$ is a Banach algebra and N is a subspace of A with the property that for all $a \in A$ and all $b \in N$, $ab \in N$, then we shall say N is a **left ideal of** A.

Lemma 15.18. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A. Then $N := \{a \in A : f(a^*a) = 0\}$ is a left ideal of A.*

Proof. From before we know that N is a subspace of A. Further if $a \in A$ and $b \in N$, then $f((ab)^*(ab)) = f(b^*a^*ab) = f(b^*(a^*ab)) = [a^*ab, b]_f = 0$, by the Cauchy-Schwarz inequality, and so $ab \in N$. $\square$

We shall need the following general result from linear algebra.

Lemma 15.19 (Factorisation Lemma)**.** *Suppose U, V and W are vector spaces, $g : U \to W$ is a surjective linear map, $f : U \to V$ is a linear map and $ker(g) \subseteq ker(f)$. Then there exists a linear map $h : W \to V$ such that $f = h \circ g$.*

Proof. For $y \in W$, as g is surjective, there exists an $x \in U$ with $g(x) = y$. Define $h(y) := f(x)$. If $g(x_1) = y = g(x_2)$ then $g(x_1 - x_2) = 0$ so $x_1 - x_2 \in \ker(g)$. Hence $x_1 - x_2 \in \ker(f)$ and so $f(x_1) = f(x_2)$. This shows h is well defined. It is immediate from the definition of h that $f = h \circ g$. As g and f are linear it can easily be checked h is also linear. $\square$

This lemma can be generalised to many other algebraic structures such as groups and rings. However, we shall only need the above version for vector spaces.

Suppose that $(A, \|\cdot\|)$ is a C^*-algebra, $(H, \langle\cdot,\cdot\rangle)$ is a Hilbert space and $\pi : A \to B(H)$ is a $*$-homomorphism (i.e., preserves scalar multiplication, addition, multiplication and the involution). Then we say say that the pair (π, H) is a **representation** of A. If π is an isometric $*$-homomorphism, then we say that (π, H) is an **isometric representation**. Furthermore, if $(A, \|\cdot\|)$ is a unital C^*-algebra and π is a unital $*$-homomorphism (i.e., $\pi(\mathbf{1}_A)$ is the identity operator on H), then we say that (π, H) is a **unital representation**. If there exists a vector $h \in H$ such that $\text{span}\{\pi(a)(h) : a \in A\}$ is dense in H, then we say that (π, H) is a **cyclic representation** and the vector h is called a **cyclic vector** for (π, H).

Example 15.20. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and (π, H) is a unital representation of A. Let $h \in H$ and define $f : A \to \mathbb{C}$ by $f(a) := \langle\pi(a)(h), h\rangle$. Then f is a linear functional. Further, since*

$$f(a^*a) = \langle\pi(a^*a)(h), h\rangle = \langle(\pi(a)^*\pi(a))(h), h\rangle = \langle\pi(a)(h), \pi(a)(h)\rangle \geqslant 0$$

it follows that f is a positive functional on A.

The next theorem is perhaps the most important theorem in this part of the course. In some sense it gives a converse to the above example and says that all bounded positive functionals come from a representation. Recall that if V is an inner product space then there exists a Hilbert space H containing V as a dense subspace. Furthermore the space H is unique, up to a unitary map, and is called the **Hilbert space completion of** V.

Theorem 15.21 (The GNS construction)**.** *Suppose that $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive functional on A. Then there exists a unital representation (π_f, H_f), of A and an $h_f \in H_f$ such that and $h_f \in H_f$ such that (π_f, H_f) is cyclic, with cyclic vector h_f, and $f(a) = \langle \pi_f(a)(h_f), h_f\rangle$ for all $a \in A$.*

Proof. Firstly, $N := \{a \in A : f(a^*a) = 0\}$ is a subspace of A and A/N is an inner product space with inner product $\langle a+N, b+N\rangle = f(b^*a)$. Let $a \in A$ and define $h_a : A \to A/N$ by $h_a(b) = ab + N$. Then h_a is a linear map. Define $g_a : A \to A/N$ by $g_a(b) = b + N$. Then g_a is a surjective linear map. Suppose $b \in \ker(g_a)$. Then $b \in N$, so as N is a left ideal of A, $ab \in N$ and $h_a(b) = ab + N = N$ thus $b \in \ker(h_a)$. By the Factorisation Lemma there exists a linear map $\pi(a) : A/N \to A/N$ such that $h_a = \pi(a) \circ g_a$. In particular $\pi(a)(b + N) = ab + N$ for all $b + N \in A/N$. Now, by Lemma 15.17,

$$\|\pi(a)(b + N)\|^2 = \langle ab + N, ab + N\rangle = f((ab)^*ab) \leqslant f(b^*b)\|a\|^2 = \|b + N\|^2\|a\|^2.$$

So $\pi(a)$ is bounded with $\|\pi(a)\| \leqslant \|a\|$.

Let H_f be the Hilbert space completion of A/N. Then, as A/N is a dense subset of H_f and $\pi(a)$ is uniformly continuous, $\pi(a)$ extends uniquely to a bounded linear functional on H_f, say $\pi_f(a)$. Now, $\pi_f : A \to B(H_f)$ is a well defined map. Further for $a, b \in A$ and $\lambda \in \mathbb{C}$ and $c + N \in A/N$,

$$\pi_f(ab)(c + N) = abc + N = \pi_f(a)(bc + N) = \pi_f(a)\pi_f(b)(c + N),$$

$$\pi_f(a + b)(c + N) = (a + b)c + N = (ac + N) + (bc + N) = \pi_f(a)(c + N) + \pi_f(b)(c + N),$$

$$\pi_f(\lambda a)(c + N) = \lambda ac + N = \lambda(ac + N) = \lambda\pi_f(a)(c + N),$$

so

$$\pi_f(ab) = \pi_f(a)\pi_f(b),\ \pi_f(a + b) = \pi_f(a) + \pi_f(b) \text{ and } \pi_f(\lambda a) = \lambda\pi_f(a)$$

on A/N. As A/N is dense in H_f by continuity these equations hold on all of H_f. Next we show that π_f preserves that involution. To do that we need to show that $\langle \pi_f(a^*)(h), k\rangle = \langle h, \pi_f(a)(k)\rangle$ for all $h, k \in H_f$. To this end, let $a \in A$, $b + N, c + N \in A/N$ then,

$$\begin{aligned}\langle \pi_f(a^*)(b + N), c + N\rangle &= \langle a^*b + N, c + N\rangle \\ &= f(c^*a^*b) \quad \text{this is why we used } [a,b]_f = f(b^*a) \text{ rather than } f(ab^*) \\ &= f((ac)^*b) \\ &= \langle b + N, ac + N\rangle \\ &= \langle b + N, \pi_f(a)(c + N)\rangle\end{aligned}$$

As A/N is dense in H_f and by the continuity of the inner product and of $\pi_f(a)$ we have $\langle \pi_f(a^*)(h), k\rangle = \langle h, \pi_f(a)(k)\rangle$ for all $h, k \in H_f$. Hence we have $\pi_f(a^*) = \pi_f(a)^*$. This shows that π_f is a $*$-homomorphism. It follows that (π_f, H_f) is a representation of A.

Now set $h_f := \mathbf{1}_A + N$. Then,

$$\operatorname{span}\{\pi_f(a)h_f : a \in A\} = \operatorname{span}\{\pi_f(a)(\mathbf{1}_A + N) : a \in A\} = \operatorname{span}\{a + N : a \in A\} = A/N$$

is dense in H_f and so (π_f, H_f) is cyclic with cyclic vector h_f. Next, for any $a \in A$,

$$\langle \pi_f(a)(h_f), h_f\rangle = \langle \pi_f(a)(\mathbf{1}_A + N), \mathbf{1}_A + N\rangle = \langle a + N, \mathbf{1}_A + N\rangle = f(\mathbf{1}_A^* a) = f(a).$$

Finally, note that for $a + N \in A/N$,

$$\pi_f(\mathbf{1}_A)(a + N) = a + N = I(a + N)$$

where I is the identity operator on H_f. As $\pi_f(\mathbf{1}_A)$ and I are equal on a dense subset of H_f by continuity it follows $\pi_f(\mathbf{1}_A) = I$. It follows (π_f, H_f) is a unital representation. $\square$

Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra and f is a positive linear functional on A. Let (π_f, H_f) be the representation of A as constructed above. Then we call (π_f, H_f) the **GNS representation of A corresponding to f**.

We saw earlier that the proof of the commutative Gelfand-Naimark Theorem relied upon an ample supply of nonzero multiplicative linear functionals. Enough in fact that for every $a \in A$ there existed a nonzero multiplicative linear functional x^* such that $|x^*(a)| = \|a\|$. However, as the next example shows, we cannot in general, expect a large supply of multiplicative linear functions.

Example 15.22. *Consider the finite dimensional Hilbert space $\mathbb{C}^n$, endowed with the usual inner product. Then $B(\mathbb{C}^n)$ has no non-trivial ideals. In particular, there are no nonzero multiplicative linear functionals on $B(\mathbb{C}^n)$, as the kernel of such a functional would be a proper ideal in $B(\mathbb{C}^n)$.*

Proof. Suppose that J is an ideal of $B(\mathbb{C}^n)$ containing a nonzero operator $A \in J$. Then there is at least one vector $z \in \mathbb{C}^n$ such that $A(z) \neq 0$. For each $1 \leqslant i \leqslant n$, let $B_i \in B(\mathbb{C}^n)$ and $C_i \in B(\mathbb{C}^n)$ be defined by, $B_i(x) := \langle x, e_i\rangle z$ and $C_i(x) = \langle x, A(z)\rangle e_i$. Let $x \in \mathbb{C}^n$, then

$$(C_i A B_i)(x) = C_i A(B(x)) = C_i A(\langle x, e_i\rangle z) = \langle x, e_i\rangle C_i(A(x)) = \langle x, e_i\rangle \|A(z)\|^2 e_i.$$

Let $D_i := \frac{1}{\|A(z)\|^2} C_i A B_i \in J$, then $D_i(x) = \langle x, e_i\rangle e_i$ and so for each $x \in \mathbb{C}^n$,

$$(\textstyle\sum_{i=1}^n D_i)(x) = \sum_{i=1}^n D_i(x) = \sum_{i=1}^n \langle x, e_i\rangle e_i = x = I_n(x).$$

Thus, $I_n = \sum_{i=1}^n D_i \in J$. This shows that $J = B(\mathbb{C}^n)$. $\square$

The next theorem shows that there are plenty of states on a unital C^*-algebra.

Theorem 15.23. *Let a be any normal element of a unital C^*-algebra $(A, \|\cdot\|)$ and let $\lambda \in \sigma_A(a)$. Then there exists a state $f \in A^*$ such that $f(a) = \lambda$. In particular, since $\|a\| = r_A(a)$ there exists a state $f \in A^*$ such that $|f(a)| = \|a\|$.*

Proof. Consider $C(a, \mathbf{1}_A)$. This is a commutative unital C^*-algebra and hence the Gelfand transform is an isomorphism from $C(a, \mathbf{1}_A)$ onto $C_{\mathbb{C}}(\Delta_{C(a)})$. Since

$$\lambda \in \sigma_A(a) = \sigma_{C(a,\mathbf{1}_A)}(a) = \mathrm{range}(\widehat{a})$$

there exists an $x^* \in \Delta_{C(a,\mathbf{1}_A)}$ such that $\lambda = \widehat{a}(x^*) = x^*(a)$. By the Hahn-Banach extension theorem there exists an $f \in A^*$ such that $\|f\| = \|x^*\| = 1$ and $f|_{C(a,\mathbf{1}_A)} = x^*$. In particular, since x^* is a nonzero multiplicative linear functional $f(\mathbf{1}_A) = x^*(\mathbf{1}_A) = 1$. Thus, f is a state and $f(a) = x^*(a) = \lambda$. $\square$

Corollary 15.24. *Let a be any element of a unital C^*-algebra $(A, \|\cdot\|)$. Then there exists a state $f \in A^*$ such that $f(a^*a) = \|a\|^2$.*

Proof. Since a^*a is self-adjoint it is normal. Therefore, by Theorem 15.23, there exists a state such that $|f(a^*a)| = \|a^*a\| = \|a\|^2$. However, as all states are positive functionals, $f(a^*a) \in [0, \infty)$. Therefore, $f(a^*a) = \|a\|^2$. $\square$

In the next section we will prove the following theorem.

Theorem 15.25 (Gelfand-Naimark, 1943)**.** *Suppose $(A, \|\cdot\|)$ is a C^*-algebra. Then there exists a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ such that $(A, \|\cdot\|)$ is isometrically $*$-isomorphic to a C^*-subalgebra of $B(H)$.*

Chapter 16

Gelfand-Naimark Theorem

Lemma 16.1. *Suppose that $(A, \|\cdot\|)$ and $(B, \|\!|\cdot|\!\|)$ are C^*-algebras and $\pi : A \to B$ is an isometric unital $*$-homomorphism. Then $\pi(A)$ is a C^*-subalgebra of B and A is isometrically $*$-isomorphic to $\pi(A)$.*

Proof. As $\pi : A \to B$ is an isometric unital $*$-homomorphism it follows that $\pi(A)$ is closed in the norm topology. It is easy to see that $\pi(A)$ is closed under multiplication and the involution. It then follows that $\pi(A)$ is a C^*-subalgebra of B that is isometrically $*$-isomorphic to A. □

As a corollary of the above result, to prove the Gelfand-Naimark Theorem it suffices to show every C^*-algebra has an isometric representation.

Let Λ be a nonempty set and for each $\lambda \in \Lambda$, let $(H_\lambda, \langle\cdot,\cdot\rangle_\lambda)$ be a Hilbert space. Note that for each $\lambda \in \Gamma$, $\|\cdot\|_\lambda^2 = \langle\cdot,\cdot\rangle_\lambda$ We define,

$$\bigoplus_{\lambda\in\Lambda} H_\lambda := \left\{ (h_\lambda)_{\lambda\in\Lambda} \in \prod_{\lambda\in\Lambda} H_\lambda : \sum_{\lambda\in\Lambda} \|h_\lambda\|_\lambda^2 < \infty \right\}.$$

If scalar multiplication and addition are defined pointwise, that is,

$$\alpha(h_\lambda)_{\lambda\in\Lambda} + \beta(k_\lambda)_{\lambda\in\Lambda} = (\alpha h_\lambda + \beta k_\lambda)_{\lambda\in\Lambda}$$

and

$$\langle (h_\lambda)_{\lambda\in\Lambda}, (k_\lambda)_{\lambda\in\Lambda} \rangle := \sum_{\lambda\in\Lambda} \langle h_\lambda, k_\lambda\rangle_\lambda,$$

then $\bigoplus_{\lambda\in\Lambda} H_\lambda$ is a Hilbert space with inner product $\langle\cdot,\cdot\rangle$.

Further, if for each $\lambda \in \Lambda$, T_λ is a bounded linear operator on H_λ and $\sup_{\lambda\in\Lambda}\|T_\lambda\| < \infty$, then

$$\bigoplus_{\lambda\in\Lambda} T_\lambda((h_\lambda)_{\lambda\in\Lambda}) := (T_\lambda(h_\lambda))_{\lambda\in\Lambda}$$

defines a bounded linear operator on $\bigoplus_{\lambda\in\Lambda} H_\lambda$ with $\|\bigoplus_{\lambda\in\Lambda} T_\lambda\| = \sup_{\lambda\in\Lambda}\|T_\lambda\|$. The proofs of these claims are straightforward calculations.

Lemma 16.2. *Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra, Λ is a nonempty set and for each $\lambda \in \Lambda$, (π_λ, H_λ) is a unital representation of A. Set $H := \bigoplus_{\lambda\in\Lambda} H_\lambda$ and define $\pi : A \to B(H)$ by $\pi(a) := \bigoplus_{\lambda\in\Lambda} \pi_\lambda(a)$. Then,*

(i) (π, H) is a unital representation of A.

(ii) If for each $a \in A\backslash\{0\}$ there exists $\lambda \in \Lambda$ with $\|\pi_\lambda(a)\| = \|a\|$, then (π, H) is isometric.

Proof. Since π_λ is a unital $*$-homomorphism, $\|\pi_\lambda(a)\| \leqslant \|a\|$ for all $\lambda \in \Lambda$. Therefore, $\sup_\lambda\|\pi_\lambda(a)\| \leqslant \|a\|$ so $\pi(a) = \bigoplus_{\lambda\in\Lambda} \pi_\lambda(a)$ is a bounded linear operator for each $a \in A$. Some straightforward calculations show that π is a unital $*$-homomorphism and so (π, H) is a unital representation of A.

Suppose that for each $a \in A\backslash\{0\}$ there exists a $\lambda \in \Lambda$ with $\|\pi_\lambda(a)\| = \|a\|$. As π is a $*$-homomorphism, π is norm decreasing and so it follows that

$$\|a\| \geqslant \|\pi(a)\| = \|\bigoplus_{\lambda\in\Lambda} \pi_\lambda(a)\| = \sup_{\lambda\in\Lambda}\|\pi_\lambda(a)\| = \|a\|.$$

This completes the proof. □

Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra, Λ is a nonempty set and for each $\lambda \in \Lambda$, (π_λ, H_λ) is a unital representation of A. Further, suppose that H and $\pi : A \to B(H)$ are defined as above, then we say that (π, H) is the **direct sum of** $((\pi_\lambda, H_\lambda))_{\lambda\in\Lambda}$.

Suppose $(A, \|\cdot\|)$ is a unital C^*-algebra. For each $f \in S(A)$, let (π_f, H_f) be the GNS representation corresponding to f with cyclic vector h_f. Let (π, H) be the direct sum of $((\pi_f, H_f))_{f\in S(A)}$. We shall call (π, H) the **universal representation of** A.

Lemma 16.3. *Suppose that $(A, \|\cdot\|)$ is a unital C^*-algebra and (π, H) is a unital representation of A. Let $h \in H$ and define $f(a) = \langle\pi(a)(h), h\rangle$ for each $a \in A$. Then $\|f\| = \|h\|^2$.*

Proof. From earlier we know that f is a positive linear functional. Using the Cauchy-Schwarz inequality and the fact that unital $*$-homomorphisms are norm decreasing we get that,

$$|f(a)| = |\langle\pi(a)(h), h\rangle| \leqslant \|\pi(a)(h)\|\|h\| \leqslant \|\pi(a)\|\|h\|^2 \leqslant \|a\|\|h\|^2$$

so $\|f\| \leqslant \|h\|^2$. Further, as $f(\mathbf{1}_A) = \langle\pi(\mathbf{1}_A)(h), h\rangle = \langle h, h\rangle = \|h\|^2$ it follows $\|f\| = \|h\|^2$. This completes the proof. □

Theorem 16.4 (Gelfand-Naimark Theorem)**.** *Suppose that $(A, \|\cdot\|)$ is a unital C^*-algebra. Then there exists an isometric unital representation of A.*

Proof. Let (π, H) be the universal representation of A. Then (π, H) is a direct sum of unital representations and so is itself a unital representation. It remains to show (π, H) is isometric. For each $a \in A\backslash\{0\}$, there exists a state, $f \in S(A)$ such that $f(a^*a) = \|a\|^2$. Let (π_f, H_f) be the GNS representation corresponding to f with cyclic vector h_f. Then as $f(a) = \langle\pi_f(a)(h_f), h_f\rangle$ we have,

$$1 = \|f\| = \|h_f\|^2$$

by Lemma 16.3. Furthermore,

$$\begin{aligned}
\|a\|^2 &= f(a^*a) \\
&= \langle \pi_f(a^*a)(h_f), h_f \rangle \\
&= \langle \pi_f(a)^*\pi_f(a)(h_f), h_f \rangle \\
&= \langle \pi_f(a)(h_f), \pi_f(a)(h_f) \rangle \\
&= \|\pi_f(a)(h_f)\|^2 \\
&\leqslant \|\pi_f(a)\|^2 \|h_f\|^2 \\
&= \|\pi_f(a)\|^2 \\
&\leqslant \|a\|^2,
\end{aligned}$$

where we used the fact that unital $*$-homomorphisms are norm decreasing. Equality is forced in the middle and so $\|\pi_f(a)\| = \|a\|$. From our earlier results, it follows that (π, f) is an isometric representation. $\square$

The question now remains as to how we handle non-unital C^*-algebras.

Unitisation

In Gelfand and Naimark's 1943 paper the C^*-algebras were assumed to be unital among other conditions. Later on it became apparent that this excluded many interesting examples such as the space of compact operators on an infinite-dimensional Hilbert space. Nevertheless, C^*-algebras with a unit are easier to work with. The aim of this section is to describe how to appropriately embed a non-unital C^*-algebra inside a unital C^*-algebra. This enables many results to be proved assuming a multiplicative identity and then extending to the non-unital case.

Lemma 16.5. *Suppose that $(X, \|\cdot\|)$ is a Banach space, S is a closed subspace of $(X, \|\cdot\|)$ and T is a finite dimensional subspace of $(X, \|\cdot\|)$. Then $S + T$ is a closed subspace of $(X, \|\cdot\|)$.*

Proof. It is easy to see $S + T$ is a subspace of X. As S is closed X/S is a Banach space with norm $\|x + S\| = \operatorname{dist}(x, S)$. Let $\pi : X \to X/S$ be the quotient map. Then π is a linear map and as T is finite dimensional, $\pi(T)$ is finite dimensional and hence closed. Therefore $S + T = \pi^{-1}(\pi(T))$ is the inverse image of a closed set so is closed. $\square$

Suppose that $(A, \|\cdot\|)$ is a C^*-algebra. For $a \in A$ define $L_a : A \to A$ by $L_a(b) := ab$ for all $b \in A$. Let I denote the identity operator in $B(A)$. For $a \in A$ and $\lambda \in \mathbb{C}$ define $L_{(a,\lambda)} := L_a + \lambda I$. Let $L_A := \{L_a : a \in A\} \subseteq B(A)$ and $L_{A\times\mathbb{C}} := \{L_a + \lambda I : a \in A, \lambda \in \mathbb{C}\} \subseteq B(A)$.

Lemma 16.6. *Suppose that $(A, \|\cdot\|)$ is a C^*-algebra. Then,*

(i) $\|L_a\| = \|a\|$ *for all* $a \in A$.

(ii) L_A *is a closed subspace of* $B(A)$.

(iii) *$L_{A\times\mathbb{C}}$ is a closed subspace of $B(A)$.*

Proof. For $b \in A$, $\|L_a(b)\| = \|ab\| \leqslant \|a\|\|b\|$ so L_a is bounded and $\|L_a\| \leqslant \|a\|$. It is easy to see L_A is a subspace of $B(A)$. Moreover,

$$\|a\|^2 = \|aa^*\| = \|L_a(a^*)\| \leqslant \|L_a\|\|a^*\| = \|L_a\|\|a\|.$$

This shows $\|L_a\| = \|a\|$. The map $A \to L_A : a \to L_a$ is surjective and isometric. L_A is the isometric image of a complete space and so is closed. Further, $\mathbb{C}I := \{\lambda I : \lambda \in \mathbb{C}\}$ is a finite dimensional subspace of $B(A)$. It follows $L_{A\times\mathbb{C}} = L_A + \mathbb{C}I$ is a closed subspace of $B(A)$, by Lemma 16.5. $\square$

Lemma 16.7. *Suppose that $(A, \|\cdot\|)$ is a Banach algebra that is also a $*$-algebra and satisfies $\|a\|^2 \leqslant \|a^*a\|$ for all $a \in A$. Then $(A, \|\cdot\|)$ is a C^*-algebra.*

Proof. Let $a \in A$. Then $\|a\|^2 \leqslant \|a^*a\| \leqslant \|a^*\|\|a\|$, so $\|a\| \leqslant \|a^*\|$. By considering a^* we also have $\|a^*\| \leqslant \|a\|$ and so $\|a^*\| = \|a\|$. Then $\|a\|^2 \leqslant \|a^*a\| \leqslant \|a^*\|\|a\| = \|a\|^2$ and so $\|a\|^2 = \|a^*a\|$. $\square$

Lemma 16.8. *Suppose that A is a $*$-algebra and $b \in A$ is a left identity for A. Then b is also a right identity for A, and so b is an identity for A.*

Proof. As b is a left identity for A, $ba = a$ for all $a \in A$. Then by taking the involution of this we get that b^* is a right identity for A. Therefore, $b = bb^* = b^*$ and so b is both a left and right identity for A, and hence an identity for A. $\square$

Define $\tilde{A} := A\times\mathbb{C}$. On $\tilde{A}$ we may define scalar multiplication $\alpha(a,\lambda) := (\alpha a, \alpha\lambda)$, addition $(a_1,\lambda_1) + (a_2,\lambda_2) := (a_1 + a_2, \lambda_1 + \lambda_2)$, an involution $(a,\lambda)^* := (a^*, \overline{\lambda})$ and multiplication $(a_1,\lambda_1)(a_2,\lambda_2) := (a_1a_2 + \lambda_1a_2 + \lambda_2a_1, \lambda_1\lambda_2)$.

Then one can check that with these operations $\tilde{A}$ is a unital $*$-algebra over $\mathbb{C}$ with multiplicative identity $(0,1)$. Let $\pi : \tilde{A} \to B(A)$ be defined by $\pi((a,\lambda)) = L_{(a,\lambda)}$.

Lemma 16.9. *Suppose that $(A, \|\cdot\|)$ is a non-untial C^*-algebra. Then the function $\pi : \tilde{A} \to B(A)$ defined above is an injective homomorphism*

Proof. It is easy to see π is linear. Let $(a_1,\lambda_1), (a_2,\lambda_2) \in \tilde{A}$. Then for $c \in A$,

$$\begin{aligned}\pi((a_1,\lambda_1)(a_2,\lambda_2))(c) &= \pi((a_1a_2 + \lambda_1a_2 + \lambda_2a_1, \lambda_1\lambda_2))(c)\\ &= L_{(a_1a_2+\lambda_1a_2+\lambda_2a_1,\lambda_1\lambda_2)}(c)\\ &= a_1a_2c + \lambda_1a_2c + \lambda_2a_1c + \lambda_1\lambda_2c\\ &= L_{(a_1,\lambda_1)}(a_2c + \lambda_2c)\\ &= L_{(a_1,\lambda_1)}L_{(a_2,\lambda_2)}(c).\end{aligned}$$

This shows π preserves multiplication and so is a homomorphism. Suppose $\pi((a,\lambda)) = L_{(a,\lambda)} = 0$. If $\lambda = 0$, then $0 = \|L_{(a,\lambda)}\| = \|L_a\| = \|a\|$ and so $a = 0$ also. If $\lambda \neq 0$, then $ac + \lambda c = 0$ for all $c \in A$, so $\frac{-1}{\lambda}a$ is a left identity for A and so is an identity for A by Lemma 16.8 which contradicts A being non-unital. It follows $(a,\lambda) = 0$ and so π is an injective homomorphism. $\square$

If A is non-unital, then from Lemma 16.9 it follows that $L_{A\times\mathbb{C}}$ is a subalgebra of $B(A)$ and that π is an isomorphism from $\tilde{A}$ onto $L_{A\times\mathbb{C}}$. Define $|||\cdot|||$ on $\tilde{A}$ by

$$|||(a,\lambda)||| := \|\pi((a,\lambda))\| = \|L_{(a,\lambda)}\|.$$

As π is an isometric isomorphism and $L_{A\times\mathbb{C}}$ is a closed subspace of the complete space $B(A)$, it follows that $(\tilde{A}, |||\cdot|||)$ is a Banach algebra.

Theorem 16.10. *Suppose that $(A, \|\cdot\|)$ is a non-unital C^*-algebra. Then $(\tilde{A}, |||\cdot|||)$ is a unital C^*-algebra.*

Proof. $(\tilde{A}, |||\cdot|||)$ is a unital Banach algebra that is also a $*$-algebra. It remains to verify the C^*-condition. Let $(a,\lambda) \in \tilde{A}$. Then,

$$\begin{aligned}
|||(a,\lambda)|||^2 &= \|L_{(a,\lambda)}\|^2 \\
&= \sup_{\|b\|\leqslant 1} \|ab+\lambda b\|^2 \\
&= \sup_{\|b\|\leqslant 1} \|(ab+\lambda b)^*(ab+\lambda b)\| \\
&= \sup_{\|b\|\leqslant 1} \|b^*(a^*ab+\lambda a^*b+\overline{\lambda}ab+|\lambda|^2 b)\| \\
&\leqslant \sup_{\|b\|\leqslant 1} \|b^*\|\|a^*ab+\lambda a^*b+\overline{\lambda}ab+|\lambda|^2 b\| \\
&\leqslant \sup_{\|b\|\leqslant 1} \|a^*ab+\lambda a^*b+\overline{\lambda}ab+|\lambda|^2 b\| \\
&= \|L_{(a^*a+\lambda a^*+\overline{\lambda}a, |\lambda|^2)}\| \\
&= \|\pi((a^*a+\lambda a^*+\overline{\lambda}a, |\lambda|^2))\| \\
&= |||(a^*a+\lambda a^*+\overline{\lambda}a, |\lambda|^2)||| \\
&= |||(a,\lambda)^*(a,\lambda)|||.
\end{aligned}$$

By Lemma 16.7, it follows that the C^*-condition is satisfied and so $L_{A\times\mathbb{C}}$ is a unital C^*-algebra. $\square$

Lemma 16.11. *If $(A, \|\cdot\|)$ is a non-unital C^*-algebra, then $A\times\{0\}$ is a C^*-subalgebra of $(\tilde{A}, |||\cdot|||)$ and the map $i_A : A \to A\times\{0\}$ defined by $i_A(a) := (a,0)$ is an isometric $*$-isomorphism.*

Proof. It is easy to check that $A\times\{0\}$ is a subspace of $\tilde{A}$ that is closed under multiplication and the involution. It is also easy to check that i_A is a linear map which preserves multiplication and the involution. By Lemma 16.6,

$$|||i_A(a)||| = |||(a,0)||| = \|\pi((a,0))\| = \|L_{(a,0)}\| = \|L_a\| = \|a\|.$$

So i_A is isometric and $A\times\{0\}$ is the isometric image of the complete space A, and so is closed. It follows that $A\times\{0\}$ is a C^*-subalgebra of $\tilde{A}$. Obviously, i_A is surjective, so it follows that i_A is an isometric $*$-isomorphism. $\square$

If $(A, \|\cdot\|)$ is a non-unital C^*-algebra, then we call $(\tilde{A}, \|\!|\cdot|\!\|)$ the **unitisation** of $(A, \|\cdot\|)$.

We can now present the full Gelfand-Naimark Representation Theorem.

Theorem 16.12 (Gelfand-Naimark Theorem*)**.** *Suppose that $(A, \|\cdot\|)$ is a C^*-algebra. Then there exists an isometric representation of A.*

Proof. If $(A, \|\cdot\|)$ is non-unital consider $(\tilde{A}, \|\!|\cdot|\!\|)$, the unitisation of $(A, \|\cdot\|)$. Let (π, H) be the universal representation of $(\tilde{A}, \|\!|\cdot|\!\|)$. The inclusion $*$-homomorphism $i_A : A \to \tilde{A}$ is an isometric $*$-homomorphism. The composition of two isometric $*$-homomorphisms is again an isometric $*$-homomorphism. Therefore, $(\pi \circ i_A, H)$ is an isometric representation of A. $\square$

We have shown that every C^*-algebra has an isometric representation. Hence, by Lemma 16.1, we also have the following version of the Gelfand-Naimark Theorem.

Theorem 16.13 (Gelfand-Naimark Theorem**)**.** *Suppose $(A, \|\cdot\|)$ is a C^*-algebra. Then there exists a Hilbert space, $(H, \langle\cdot,\cdot\rangle)$, such that A is isometrically $*$-isomorphic to a C^*-subalgebra of $B(H)$.*

Chapter 17

Compact Operators

Let $(X, \|\cdot\|)$ and $(Y, \|\!|\cdot, \|\!|)$ be normed linear spaces and let $T \in B(X, Y)$. Then T is called a **compact operator** if $\overline{T(B_X)}$ is a compact subset of $(Y, \|\!|\cdot, \|\!|)$. Clearly if either X or Y is finite dimensional, then T is a compact operator. In this section we will show that if $(X, \|\cdot\|)$ is a Banach space and $T \in B(X)$ is compact, then $I_X - T$ is 1-to-1 if, and only if, $I_X - T$ is onto. Moreover, both of these conditions are equivalent to $I_X - T$ being an isomorphism from $(X, \|\cdot\|)$ onto $(X, \|\cdot\|)$.

Theorem 17.1. *Given a compact operator T on a Banach space $(X, \|\cdot\|)$, if $I_X - T$ is 1-to-1, then $I_X - T$ has a continuous inverse on $(I_X - T)(X)$. In particular, $(I_X - T)(X)$ is a closed subspace of $(X, \|\cdot\|)$.*

Proof. Let $m := \inf\{\|(I_X - T)(x)\| : x \in S_X\}$. Claim: $m > 0$. To prove this let us suppose, in order to obtain a contradiction, that $m = 0$. Then there exists a sequence $(x_n : n \in \mathbb{N})$ in S_X such that $\lim_{n\to\infty} \|(I_X - T)(x_n)\| = 0$. Since

$$\{T(x_n) : n \in \mathbb{N}\} \subseteq \overline{T(B_X)},$$

$(x_n : n \in \mathbb{N})$ possesses a subsequence $(x_{n_k} : k \in \mathbb{N})$ such that $y := \lim_{k\to\infty} T(x_{n_k})$. Then,

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} I_X(x_{n_k}) = \lim_{k\to\infty} (I_X - T)(x_{n_k}) + \lim_{k\to\infty} T(x_{n_k}) = 0 + y = y$$

and so $y \in S_X$, since $\{x_{n_k} : k \in \mathbb{N}\} \subseteq S_X$. On the other hand,

$$\|(I_X - T)(y)\| = \left\|(I_X - T)\left(\lim_{k\to\infty} x_{n_k}\right)\right\| = \lim_{k\to\infty} \|(I_X - T)(x_{n_k})\| = 0.$$

Therefore, $y \in \mathrm{Ker}(I_X - T) \cap S_X = \varnothing$, as $\mathrm{Ker}(I_X - T) = \{0\}$. Hence, we have obtained our desired contradiction and so it must be the case that $m > 0$.

Now, $\|(I_X - T)(x)\| \geqslant m\|x\|$ for all $x \in X$ and so $(I_X - T)$ is an isomorphism onto $(I_X - T)(X)$. For the justification for this, see the first "fact" in the chapter on conjugate mappings. $\square$

Let $(X, \|\cdot\|)$ be a Banach space. We shall denote by $\mathcal{K}(X)$ the set of all compact operators on $(X, \|\cdot\|)$. It is easy to show that $\mathcal{K}(X)$ is an **ideal** in $B(X)$, that is, (i) $\mathcal{K}(X)$ is a vector subspace of $B(X)$; (ii) $T \circ S \in \mathcal{K}(X)$ for all $T \in B(X)$ and all $S \in \mathcal{K}(X)$ and $S \circ T \in \mathcal{K}(X)$ for all $S \in \mathcal{K}(X)$ and all $T \in B(X)$. Note that: (i) $I_X \in \mathcal{K}(X)$ if, and only if, X is finite dimensional and (ii) if $(H, \langle\cdot,\cdot\rangle)$ is a Hilbert space, then $\mathcal{K}(X)$ is closed under the adjoint operation on $B(H)$. This follows from the original definition of the adjoint operation and Schauder's Theorem, see Exercise 11.6.

Lemma 17.2. *Given a compact operator T on a Banach space $(X, \|\cdot\|)$, for each $n \in \mathbb{N}$, $(I_X - T)^n = I_X - S_n$, where, S_n is a compact operator on $(X, \|\cdot\|)$.*

Proof. Let $n \in \mathbb{N}$, then

$$(I_X - T)^n = \sum_{j=0}^{n} (-1)^j \binom{n}{j} T^j = I_X - S_n, \text{ where, } S_n := \sum_{j=1}^{n} (-1)^{(j-1)} \binom{n}{j} T^j.$$

Now, S_n is compact, since $\mathcal{K}(X)$ is an ideal in $B(X)$. □

Theorem 17.3. *Given a compact operator T on a Banach space $(X, \|\cdot\|)$, if $I_X - T$ is 1-to-1, then $I_X - T$ is onto, and so an isomorphism on $(X, \|\cdot\|)$.*

Proof. For each $n \in \mathbb{N}$, let $X_n := (I_X - T)^n(X)$. Then clearly,

$$\cdots \subseteq X_{n+1} \subseteq X_n \subseteq \cdots X_2 \subseteq X_1 \subseteq X.$$

Suppose that X_{n+1} is a proper subspace of X_n for all $n \in \mathbb{N}$. By Theorem 17.1, X_{n+1} is a closed subspace of X_n, so by Riesz's Lemma (Lemma 2.16) there exists an $x_n \in S_{X_n}$ such that $\operatorname{dist}(x_n, X_{n+1}) > 1/2$. Now, for any $n > m$ we have

$$\|T(x_m) - T(x_n)\| = \|x_m - [(I_X - T)(x_m) + x_n - (I_X - T)(x_n)]\| > 1/2$$

since $[(I_X - T)(x_m) + x_n - (I_X - T)(x_n)] \in X_{m+1}$. Also, $\{x_n : n \in \mathbb{N}\} \subseteq B_X$, but $(T(x_n) : n \in \mathbb{N})$ has no convergent subsequences; which is impossible since T is a compact operator. Hence, there must be some $m \in \mathbb{N}$ such that $X_{m+1} = X_m$.

Since $(I_X - T)$ is 1-to-1, $(I_X - T)^m$ is 1-to-1. Now let x be any element of X. Then $(I_X - T)^m(x) \in X_m = X_{m+1} = (I_X - T)^{m+1}(X)$ and so there is some $y \in X$ such that

$$(I_X - T)^m(x) = (I_X - T)^{m+1}(y) = (I_X - T)^m((I_X - T)(y)).$$

However, since $(I_X - T)^m$ is 1-to-1, $x = (I_X - T)(y) \in X_1$. Therefore, $X_1 = X$, i.e., $(I_X - T)$ is onto. The fact that $I_X - T$ is an isomorphism now follows from the Open Mapping Theorem, see Theorem 6.2. □

Theorem 17.4. *Given a compact operator T on a Banach space $(X, \|\cdot\|)$, if $I_X - T$ is onto, then $I_X - T$ is 1-to-1, and so an isomorphism on $(X, \|\cdot\|)$.*

Proof. Since $I_X - T$ is onto, its conjugate $(I_X - T)' = I_{X^*} - T'$ is 1-to-1 on X^*, see the second "fact" in the conjugate mapping chapter. Since T is compact, by Schauder's Theorem, its conjugate T' is also compact. It then follows from Theorem 17.3 that $I_{X^*} - T'$ is an isomorphism on $(X^*, \|\cdot\|)$, and so from Theorem 8.4, $I_X - T$ is an isomorphism on $(X, \|\cdot\|)$. $\square$

Corollary 17.5 (Fredholm Alternative). *For a compact operator on a Banach space $(X, \|\cdot\|)$ the following are equivalent: (i) $I_X - T$ is 1-to-1; (ii) $I_X - T$ is onto; (iii) $I_X - T$ is an isomorphism on $(X, \|\cdot\|)$.*

Given a linear operator T on a vector space X, over $\mathbb{K}$, an **eigenvalue** of T is element λ of $\mathbb{K}$ such that $T(x) = \lambda x$ for some nonzero vector $x \in X$, i.e., λ is an eigenvalue of T if $\operatorname{Ker}(T - \lambda I_X) \neq \{0\}$. A nonzero vector $x \in X$ is called an **eigenvector** of T if there exists an element $\lambda \in \mathbb{K}$ such that $T(x) = \lambda x$. The **eigenspace** corresponding to an eigenvalue λ is equal to the kernel of $(T - \lambda I_X)$, i.e., it is the set of all eigenvectors (corresponding to the eigenvalue λ) plus the zero vector.

Theorem 17.6. *Let T be a compact operator defined on a Banach space $(X, \|\cdot\|)$. Then each element of $\sigma(T) \setminus \{0\}$ is an eigenvalue of T.*

Proof. Suppose that $\lambda \in \sigma(T) \setminus \{0\}$, then $T - \lambda I_X = -\lambda(I_X - \lambda^{-1}T)$ is not an isomorphism. Therefore, $I_X - \lambda^{-1}T$ is not an isomorphism. Now, as $\lambda^{-1}T$ is a compact operator, we have by Corollary 17.5, that $I_X - \lambda^{-1}T$ is not an isomorphism if, and only if, $I_X - \lambda^{-1}T$ is not 1-to-1, i.e., if, and only if, there exists a nonzero $x \in X$ such that $(I_X - \lambda^{-1}T)(x) = 0$, i.e., $T(x) = \lambda x$. $\square$

Theorem 17.7. *Let T be a compact normal operator defined on a nontrivial Hilbert space $(H, \langle\cdot,\cdot\rangle)$, over $\mathbb{C}$. Then T has an eigenvalue $\lambda \in \mathbb{C}$, such that $|\lambda| = \|T\|$.*

Proof. First, if $\|T\| = 0$, then the result is obvious. Hence we may assume that that $\|T\| > 0$. Since T is a normal operator $r(T) = \|T\| > 0$. Thus we may choose $\lambda \in \sigma(T)$ such that $|\lambda| = r(T) > 0$. Then, by Theorem 17.6, λ is an eigenvalue of T. $\square$

More facts concerning compact operators

Theorem 17.8. *For a compact operator T on a Banach space $(X, \|\cdot\|)$, Ker$(I_X - T)$ is finite dimensional. In particular, for each nonzero eigenvalue λ of T, the eigenspace corresponding to λ is finite dimensional.*

Proof. Let $Y := \operatorname{Ker}(I_X - T)$. Then Y is a closed subspace of $(X, \|\cdot\|)$. Notice that if $y \in Y$, then $T(y) = I_X(y) = y$. That is, $T|_Y = I_Y$. However, $T|_Y$ is a compact operator and so $B_Y = I_Y(B_Y) = T|_Y(B_Y) \subseteq \overline{T|_Y(B_Y)}$; which is compact. Therefore, Y is finite dimensional. $\square$

Theorem 17.9. *Let T be a linear operator defined on a vector space X. If $\{e_1, e_2, \ldots, e_n\}$ are eigenvectors of T corresponding to distinct eigenvalues of T, then $\{e_1, e_2, \ldots, e_n\}$ is a linearly independent set.*

Proof. The proof of this is left as an exercise for the reader. $\square$

Theorem 17.10. *Let T be a compact operator defined on a Banach space $(X, \|\cdot\|)$. Then $\sigma(T)$ is at most countable. Moreover, if $\sigma(T)$ has infinitely many elements, then they may be listed as a sequence that converges to* 0.

Proof. To prove the statement of the theorem it is sufficient to show that for each $\varepsilon > 0$, $\{z \in \sigma(T) : \varepsilon < |z|\}$ is finite. To this end, fix $\varepsilon > 0$ and suppose that there is an infinite sequence $(\lambda_n : n \in \mathbb{N})$ of distinct elements of $\sigma(T) \setminus \varepsilon B_{\mathbb{C}}$. For each $n \in \mathbb{N}$, let $M_n := \text{span}\{e_1, e_2, \ldots e_n\}$, where e_k is an eigenvector of T (with unit length) corresponding to the eigenvalue λ_k. Next, for each $n > 1$, choose $x_n \in S_X \cap M_n$ such that $\text{dist}(x_n, M_{n-1}) > 1/2$. Then note that $x_n - \lambda_n^{-1}T(x_n) \in M_{n-1}$, since if $x_n = \sum_{k=1}^n c_k e_k$, then

$$x_n - \lambda_n^{-1}T(x_n) = \textstyle\sum_{k=1}^n c_k(1 - \lambda_n^{-1}\lambda_k)e_k = \sum_{k=1}^{n-1} c_k(1 - \lambda_n^{-1}\lambda_k)e_k \in M_{n-1}.$$

Hence, $\text{dist}(\lambda_n^{-1}T(x_n), M_{n-1}) = \text{dist}(x_n, M_{n-1}) > 1/2$. Thus,

$$\text{dist}(T(x_n), M_{n-1}) > |\lambda_n|/2 \geqslant \varepsilon/2.$$

Notice that if $n > m$, then $\|T(x_n) - T(x_m)\| > \varepsilon/2$, since $T(x_m) \in M_m \subseteq M_{n-1}$. Now, $\{x_n : n \in \mathbb{N}\} \subseteq B_X$, but $(T(x_n) : n \in \mathbb{N})$ has no convergent subsequences; which is impossible since T is a compact operator. Therefore, $\{z \in \sigma(T) : \varepsilon < |z|\}$ is finite. $\square$

Let $(X, \|\cdot\|)$ be a Banach space and suppose that $T \in B(X)$. Then T is called a **finite rank** operator if $\dim(T(X)) < \infty$. We shall denote by $\mathcal{F}(X)$ the set of all finite rank operators defined on $(X, \|\cdot\|)$. Clearly, $\mathcal{F}(X) \subseteq \mathcal{K}(X)$ since bounded subsets of finite dimensional spaces are relatively compact.

Exercise 17.11. *Let $(X, \|\cdot\|)$ be a Banach space. Show that both $\mathcal{F}(X)$ and $\mathcal{K}(X)$ are ideals in $B(X)$.*

Theorem 17.12. *Let $(X, \|\cdot\|)$ be a Banach space. Then $\mathcal{K}(X)$ is a closed ideal in $B(X)$.*

Proof. Let $(T_n : n \in \mathbb{N})$ be a sequence in $\mathcal{K}(X)$ and suppose that $T = \lim_{n\to\infty} T_n$. We need to show that $\overline{T(B_X)}$ is compact. Since $(X, \|\cdot\|)$ is a Banach space it will be sufficient to show that for every $\varepsilon > 0$ there exists a compact set K in X such that $T(B_X) \subseteq K + \varepsilon B_X$, see Corollary 11.2. To this end, fix $\varepsilon > 0$ and choose $n \in \mathbb{N}$ such that $\|T - T_n\| < \varepsilon$. Then

$$T(B_X) \subseteq T_n(B_X) + \varepsilon B_X \subseteq \overline{T_n(B_X)} + \varepsilon B_X;$$

which completes the proof, since $\overline{T_n(B_X)}$ is compact. $\square$

Exercise 17.13. *Let K be a compact subset of a Banach space $(X, \|\cdot\|)$ and let $(T_n : n \in \mathbb{N})$ be a sequence in $B(X)$. Show that if $(T_n : n \in \mathbb{N})$ converges pointwise to $T \in B(X)$ on X, then $(T_n : n \in \mathbb{N})$ converges uniformly to T on K. Hint: Use the Uniform Boundedness Theorem.*

Theorem 17.14. *Let $(X, \|\cdot\|)$ be a Banach space with a Schauder basis. Then $\mathcal{K}(X) = \overline{\mathcal{F}(X)}$.*

Proof. It follows from the fact that (i) $\mathcal{F}(X) \subseteq \mathcal{K}(X)$ and (ii) $\mathcal{K}(X)$ is closed, see Theorem 17.12, that $\overline{\mathcal{F}(X)} \subseteq \mathcal{K}(X)$. So we need only show the reverse set-inclusion. To this end, let $T \in \mathcal{K}(X)$. Let $(e_n : n \in \mathbb{N})$ be a Schauder basis for $(X, \|\cdot\|)$ and let $(P_n : n \in \mathbb{N})$ be the canonical projections. Then for each $n \in \mathbb{N}$, $(P_n \circ T) \in \mathcal{F}(X)$. Now, $(P_n : n \in \mathbb{N})$ converge pointwise to $I_X \in B(X)$ on X and $\overline{T(B_X)}$ is compact. Therefore, by Exercise 17.13, $(P_n : n \in \mathbb{N})$ converges uniformly to I_X on $\overline{T(B_X)}$. Thus,

$$\lim_{n\to\infty} (P_n \circ T) = T \quad \text{with respect to the operator norm on } B(X).$$

This completes the proof. $\square$

Example 17.15. *Suppose that $K \in C_{\mathbb{C}}([0,1] \times [0,1])$. Then the mapping*

$$T : (L^2[0,1], \|\cdot\|_2) \to (C_{\mathbb{C}}[0,1], \|\cdot\|_\infty) \quad \textit{defined by,}$$

$$T(x)(t) := \int_{[0,1]} K(t,s)x(s)\ \mathrm{d}s \quad \textit{for all } t \in [0,1]$$

is a compact operator.

Proof : By the continuity of K and the compactness of $[0,1] \times [0,1]$, we have

$$M := \sup\{|K(t,s)| : (t,s) \in [0,1] \times [0,1]\} < \infty$$

and hence for any $x \in B_{L^2[0,1]}$ and any $t \in [0,1]$ we get

$$\begin{aligned} |T(x)(t)| &= \left|\int_{[0,1]} K(t,s)x(s)\ \mathrm{d}s\right| \leqslant \int_{[0,1]} |K(t,s)||x(s)|\ \mathrm{d}s \\ &\leqslant \left(\int_{[0,1]} |K(t,s)|^2\ \mathrm{d}s\right)^{\frac{1}{2}} \|x\|_2 \leqslant \left(\int_{[0,1]} M^2\ \mathrm{d}s\right)^{\frac{1}{2}} \|x\|_2 \leqslant M. \end{aligned}$$

Given $0 < \varepsilon$, it follows from the continuity of K and compactness of $[0,1] \times [0,1]$ that there exists a $0 < \delta$ such that if $t_1, t_2 \in [0,1]$ and $|t_1 - t_2| < \delta$, then $|K(t_1,s) - K(t_2,s)| < \varepsilon$ for all $s \in [0,1]$. Consequently, for every $x \in B_{L^2[0,1]}$ and every $t_1, t_2 \in [0,1]$ with $|t_1 - t_2| < \delta$ we have

$$\begin{aligned} |T(x)(t_1) - T(x)(t_2)| &= \left|\int_{[0,1]} K(t_1,s)x(s)\ \mathrm{d}s - \int_{[0,1]} K(t_2,s)x(s)\ \mathrm{d}s\right| \\ &\leqslant \int_{[0,1]} |K(t_1,s) - K(t_2,s)||x(s)|\ \mathrm{d}s \\ &\leqslant \left(\int_{[0,1]} |K(t_1,s) - K(t_2,s)|^2\ \mathrm{d}s\right)^{\frac{1}{2}} \|x\|_2 \\ &\leqslant \left(\int_{[0,1]} \varepsilon^2\ \mathrm{d}s\right)^{\frac{1}{2}} \|x\|_2 \leqslant \varepsilon. \end{aligned}$$

Therefore, $T(x) \in C_{\mathbb{C}}[0,1]$. In fact, we also showed that $T(B_{L^2[0,1]})$ is a uniformly bounded (by M) and an equicontinuous subset of $C_{\mathbb{C}}[0,1]$. Hence, by the Arzelà-Ascoli Theorem, $T(B_{L^2[0,1]})$ is relatively compact in $(C_{\mathbb{C}}[0,1], \|\cdot\|_\infty)$. $\square$

Remarks 17.16. *If we let* $I : (C_{\mathbb{C}}[0,1], \|\cdot\|_\infty) \to (L^2[0,1], \|\cdot\|_2)$ *be the natural inclusion map, then we see that* $I \circ T : (L^2[0,1], \|\cdot\|_2) \to (L^2[0,1], \|\cdot\|_2)$ *is a compact operator too, since* I *is a continuous linear operator. Hence* $T : (L^2[0,1], \|\cdot\|_2) \to (L_2[0,1], \|\cdot\|_2)$*, as defined above, may be directly viewed as a compact operator on* $(L^2[0,1], \|\cdot\|_2)$*.*

Exercise 17.17. *Suppose that* $K \in C_{\mathbb{C}}([a,b] \times [a,b])$*. Show that the mapping*

$$T : (L^2[a,b], \|\cdot\|_2) \to (L^2[a,b], \|\cdot\|_2) \quad \textit{defined by,}$$

$$T(x)(t) := \int_{[a,b]} K(t,s)x(s)\ \mathrm{d}s \quad \textit{for all } t \in [a,b]$$

is a compact operator.

Chapter 18

Spectral Mapping Theorem

Lemma 18.1. *For a normal operator T defined on a Hilbert space $(H, \langle \cdot, \cdot \rangle)$, λ is an eigenvalue of T if, and only if $\overline{\lambda}$ is an eigenvalue of T^*. Moreover, λ and $\overline{\lambda}$ have the same eigenspace.*

Proof. For any $x \in H$ and normal operator N of H,

$$\begin{aligned} \|N^*(x)\|^2 &= \langle N^*(x), N^*(x) \rangle = \langle NN^*(x), x \rangle = \langle N^*N(x), x \rangle \\ &= \langle N(x), N(x) \rangle = \|N(x)\|^2. \end{aligned}$$

That is, $\|N^*(x)\| = \|N(x)\|$. Therefore, $\|(T - \lambda I)(x)\| = \|(T^* - \overline{\lambda} I)(x)\|$ since $T - \lambda I$ is also a normal operator. $\square$

Theorem 18.2 (Spectral Mapping Theorem). *Let $(H, \langle \cdot, \cdot \rangle)$ be a complex infinite dimensional separable Hilbert space. If T is a compact normal operator on H, then there exists an orthonormal basis $(e_n)_{n=1}^{\infty}$ of H where each e_i is an eigenvector corresponding to an eigenvalue λ_i of T, such that for each $x \in H$ we have*

$$T(x) = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n$$

Moreover, for every $\lambda \notin \sigma(T)$ and $x \in H$ we have that

$$R(\lambda)(x) = \sum_{n=1}^{\infty} \frac{\langle x, e_n \rangle e_n}{\lambda_n - \lambda}.$$

Proof. Let $\mathcal{U}$ be a maximal (with respect to set inclusion) family of orthonormal eigenvectors of H. To prove the first part of the theorem it is sufficient to show that if $X := \overline{\text{span}}(\mathcal{U})$, then $H = X$.

Suppose, in order to obtain a contradiction that $X \neq H$. Then $X^\perp \neq \{0\}$. Next, let us show that $T|_{X^\perp} : X^\perp \to X^\perp$ and $T^*|_{X^\perp} : X^\perp \to X^\perp$. To see this, first note that both T

and T^* map X into X, since the members of $\mathcal{U}$ are eigenvectors for both T and T^*. Fix $y \in X^\perp$, then for any $x \in X$

$$\langle T(y), x\rangle = \langle y, T^*(x)\rangle = 0 \quad \text{and} \qquad \langle T^*(y), x\rangle = \langle y, T(x)\rangle = 0.$$

Therefore, $T(y) \in X^\perp$ and $T^*(y) \in X^\perp$. Moreover, it is easy to check that $T|_{X^\perp}$ is also a compact normal operator. Hence, by Theorem 17.7, $T|_{X^\perp}$ has an eigenvector $e \in X^\perp$ of unit length. But then $\mathcal{U} \cup \{e\}$ is an orthonormal family of eigenvectors which is strictly bigger than $\mathcal{U}$. However, this contradicts the maximality of $\mathcal{U}$. Hence, it must be the case that $X = H$. Now, because H is separable, we have, by Exercise 3.12 that $\mathcal{U}$ is at most countable, since $\|u - v\| = \sqrt{2}$, for each $u, v \in \mathcal{U}$ with $u \neq v$. Note also that $\mathcal{U}$ is an infinite set, because otherwise, $H = \overline{\text{span}}(\mathcal{U}) = \text{span}(\mathcal{U})$, would be finite dimensional.

Hence, we may enumerate $\mathcal{U}$ as $\{e_n : n \in \mathbb{N}\}$. For each $n \in \mathbb{N}$, let λ_n denote the eigenvalue of T, corresponding to the eigenvector e_n.

Then for any $x \in H$, $x = \lim_{n\to\infty} \sum_{k=1}^{n} \langle x, e_k\rangle e_k = \sum_{k=1}^{\infty} \langle x, e_k\rangle e_k$. Therefore,

$$\begin{aligned}
T(x) &= T\left(\lim_{n\to\infty} \sum_{k=1}^{n} \langle x, e_k\rangle e_k\right) \\
&= \lim_{n\to\infty} T\left(\sum_{k=1}^{n} \langle x, e_k\rangle e_k\right) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \langle x, e_k\rangle T(e_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \lambda_k \langle x, e_k\rangle e_k \\
&= \sum_{k=1}^{\infty} \lambda_k \langle x, e_k\rangle e_k.
\end{aligned}$$

Next, suppose $x \in H$ and $\lambda \notin \sigma(T)$ then for some $y \in H$ we have that:

$$x = (T - \lambda I)(y) = (T - \lambda I)\left(\sum_{n=1}^{\infty} \langle y, e_n\rangle e_n\right) = \sum_{n=1}^{\infty} (\lambda_n - \lambda)\langle y, e_n\rangle e_n.$$

Therefore, for each $n \in \mathbb{N}$, $\langle x, e_n\rangle = (\lambda_n - \lambda)\langle y, e_n\rangle$ and so

$$\langle y, e_n\rangle = \frac{\langle x, e_n\rangle}{\lambda_n - \lambda} \quad \text{for all } n \in \mathbb{N}.$$

On the other hand, $y = (T - \lambda I)^{-1}(x) = R(\lambda)(x)$. Therefore,

$$R(\lambda)(x) = \sum_{n=1}^{\infty} \frac{\langle x, e_n\rangle e_n}{\lambda_n - \lambda}.$$

This completes the proof $\square$

Index