MAT4450 Advanced Functional Analysis 2021

These notes were written by Ulrik Enstad the spring of 2021, and are based on notes written by Vegard Antun when Erik Bédos was lecturing MAT4450 the spring of 2017, handwritten notes by Erik Bédos from the spring of 2020, and the book *Topological vector spaces* by Narici and Beckenstein.

Contents

Chapter 1

Topological preliminaries

1.1 Neighborhood bases

Let X be a topological space and let $x \in X$. Recall that an *open neighborhood of* x is an open set in X that contains x. A *neighborhood of* x is a set that contains an open neighborhood of x. We denote by $\mathcal{N}_x$ the set of all neighborhoods of x and call it the *neighborhood filter* of x. We also denote by $\mathcal{U}_x$ the set of all open neighborhoods of x.

The following proposition, which we will take for granted, shows that one can define topology on a set starting from suitable candidates $\mathcal{N}_x$ for neighborhood filters.

Proposition 1.1.1. *Let X be a set. Suppose that we are given a topology $\mathcal{T}$ on X, and denote by $\mathcal{N}(\mathcal{T}) = (\mathcal{N}_x)_{x \in X}$ the collection of neighborhood filters associated to $\mathcal{T}$. Then the following properties are satisfied for all $x \in X$:*

(a) $X \in \mathcal{N}_x$.

(b) *If $N \in \mathcal{N}_x$ then $x \in N$.*

(c) *If $N \in \mathcal{N}_x$ and $N \subseteq N' \subseteq X$ then $N' \in \mathcal{N}_x$.*

(d) *If $N, N' \in \mathcal{N}_x$ then $N \cap N' \in \mathcal{N}_x$.*

(e) *If $N \in \mathcal{N}_x$ then there exists $U \in \mathcal{N}_x$ such that $U \subseteq N$ and $U \in \mathcal{N}_y$ for every $y \in U$.*

Conversely, if we are given a collection $\mathcal{N} = (\mathcal{N}_x)_{x \in X}$ of subsets of $\mathcal{P}(X)$ such that the above five axioms are satisfied, then

$$\mathcal{T}(\mathcal{N}) = \{U \subseteq X : U \in \mathcal{N}_x \text{ for all } x \in U\}$$

defines a topology on X.

These two procedures are mutually inverse: That is, if $\mathcal{T}$ is a topology on X then $\mathcal{T}(\mathcal{N}(\mathcal{T})) = \mathcal{T}$ and if $\mathcal{N} = (\mathcal{N}_x)_{x \in X}$ is a collection of subsets of $\mathcal{P}(X)$ satisfying the above five axioms then $\mathcal{N} = \mathcal{N}(\mathcal{T}(\mathcal{N}))$.

If X and Y are topological spaces and $f\colon X \to Y$ is a function, then the continuity of f can be expressed in terms of neighborhood filters as follows: We have that f is continuous at

x if and only if $f^{-1}(N) \in \mathcal{N}_x$ for every $N \in \mathcal{N}_{f(x)}$. An equivalent statement is that for every $N \in \mathcal{N}_{f(x)}$ there exists $N' \in \mathcal{N}_x$ such that $f(N') \subseteq N$.

A useful way to describe a topology on a set is by means of a *basis.* The following is a similar concept for neighborhood filters:

Definition 1.1.2. Let X be a topological space and let $x \in X$. A set $\mathcal{B}_x \subseteq \mathcal{N}_x$ is called a *neighborhood basis at* x if for every $N \in \mathcal{N}_x$ there exists $B \in \mathcal{B}_x$ such that $B \subseteq N$.

Example 1.1.3. (a) For any point x in a topological space X, both $\mathcal{N}_x$ and $\mathcal{U}_x$ are neighborhood bases at x.

(b) Let X be a metric space and let $x \in X$. Then the set $\mathcal{B}_x$ of all open balls

$$B_r(x) = \{y \in X : |x - y| < r\} \quad \text{for } r > 0,$$

forms a neighborhood basis at x.

Observation 1.1.4. If x is a point in a topological space X and $\mathcal{B}_x$ is a neighborhood basis at x, then the neighborhood filter $\mathcal{N}_x$ can be recovered from $\mathcal{B}_x$ as follows:

$$\mathcal{N}_x = \{N \subseteq X : \text{there exists } B \in \mathcal{B}_x \text{ such that } B \subseteq N\}.$$

If we have neighborhood bases $\mathcal{B}_x$ at every point $x \in X$, then it follows from Proposition 1.1.1 that the topology on X can be recovered as

$$\mathcal{T} = \{U \subseteq X : \text{for every } x \in U \text{ there exists } B \in \mathcal{B}_x \text{ such that } B \subseteq U\}.$$

When we define the topology induced by the metric on a metric space, we do it in terms of open balls: We declare a set N containing x to be a neighborhood of x if it contains an open ball centered at x. This only works because open balls have certain properties. The following proposition generalizes this procedure of defining a topology in terms of candidates for neighborhood bases:

Proposition 1.1.5. *Let X be a set. Suppose we are given a topology $\mathcal{T}$ on X and for every $x \in X$ a neighborhood basis $\mathcal{B}_x$ at x. Then the following hold for every $x \in X$:*

(a) $\mathcal{B}_x$ is nonempty.

(b) If $B \in \mathcal{B}_x$ then $x \in B$.

(c) If $B, B' \in \mathcal{B}_x$ then there exists $B'' \in \mathcal{B}_x$ such that $B'' \subseteq B \cap B'$.

(d) If $B \in \mathcal{B}_x$ then there exists a set $U \subseteq B$ containing x with the following property: For every $y \in U$ there exists $B' \in \mathcal{B}_y$ such that $B' \subseteq U$.

Conversely, suppose we have for every $x \in X$ a collection $\mathcal{B}_x$ of subsets of X that satisfies the above axioms. Then there exists a unique topology on X such $\mathcal{B}_x$ is a neighborhood basis at x with respect to the topology for every $x \in X$.

Proof. We begin by assuming a topology on $\mathcal{T}$ with neighborhood bases $\mathcal{B}_x$ at every $x \in X$. We need to show that the above axioms hold, so let $x \in X$.

(a): Since $X \in \mathcal{N}_x$ we have that there exists some $B \in \mathcal{B}_x$ with $B \subseteq X$. Hence $\mathcal{B}_x$ is nonempty.

(b): Follows from the fact that $\mathcal{B}_x \subseteq \mathcal{N}_x$.

(c): Suppose that $B, B' \in \mathcal{B}_x$. Then since B and B' are neighborhoods of x, $B \cap B'$ must be a neighborhood of x. Thus, by definition of neighborhood basis, there exists $B'' \in \mathcal{B}_x$ such that $B'' \subseteq B \cap B'$.

(d): Let $B \in \mathcal{B}_x$. Then B contains an open neighborhood U of x. If $y \in U$ then U is a neighborhood of y so there exists $B' \in \mathcal{B}_y$ such that $B' \subseteq U$.

Now, suppose we are given collections $\mathcal{B}_x$ that satisfy the above axioms. Define

$$\mathcal{N}_x = \{N \subseteq X : \text{there exists } B \in \mathcal{B}_x \text{ such that } B \subseteq N\}.$$

We will show that the sets $(\mathcal{N}_x)_{x \in X}$ satisfy the properties in Proposition 1.1.1. Let us call them $(a') - (e')$ to distinguish them from the properties in the current proposition. Let $x \in X$.

(a'): Since $\mathcal{B}_x \neq \emptyset$ by (a), there exists some $B \in \mathcal{B}_x$. But then $B \subseteq X$ so $X \in \mathcal{N}_x$.

(b'): Let $N \in \mathcal{N}_x$. Then there exists $B \in \mathcal{B}_x$ with $B \subseteq N$. Since $x \in B$ by (b) we have that $x \in N$.

(c'): Let $N \in \mathcal{N}_x$ and $N \subseteq N' \subseteq X$. Then there exists $B \in \mathcal{B}_x$ such that $B \subseteq N$. But then $B \subseteq N'$ as well, so $N' \in \mathcal{N}_x$.

(d'): Let $N, N' \in \mathcal{N}_x$. Then we can find $B, B' \in \mathcal{B}_x$ with $B \subseteq N$ and $B' \subseteq N'$. By assumption we can find $B'' \in \mathcal{B}_x$ such that $B'' \subseteq B \cap B'$. But then $B'' \subseteq N \cap N'$, so $N \cap N' \in \mathcal{N}_x$.

(e') Let $N \in \mathcal{N}_x$. Then there exists some $B \in \mathcal{B}_x$ with $B \subseteq N$. By assumption we can find $U \subseteq B$ containing x such that for all $y \in U$ there exists $B' \in \mathcal{B}_y$ with $B' \subseteq U$. Then $U \subseteq N$. Now if $y \in U$ then we can find $B' \in \mathcal{B}_y$ such that $B' \subseteq U$, i.e. $U \in \mathcal{N}_y$. In particular this applies to $y = x$, so $U \in \mathcal{N}_x$.

We have now verified the axioms of Proposition 1.1.1, so we can conclude that the sets $(\mathcal{N}_x)_{x \in X}$ are the neighborhood filters for a topology on X. It follows straight from the definition of $\mathcal{N}_x$ that $\mathcal{B}_x$ is a neighborhood basis at x. From Observation 1.1.4 we see that this is the unique topology on X with $\mathcal{B}_x$ as neighborhood basis at x for each $x \in X$. ∎

Recall that if $\mathcal{T}_1$ and $\mathcal{T}_2$ are topologies on a set X, we say that $\mathcal{T}_1$ is *finer* or *stronger* than $\mathcal{T}_2$ if $\mathcal{T}_1 \supseteq \mathcal{T}_2$. We also say that $\mathcal{T}_2$ is *coarser* or *weaker* than $\mathcal{T}_1$.

Note that given any collection $\mathcal{B}$ of subsets of X, there exists a weakest topology $\mathcal{T}$ on X such that $\mathcal{B} \subseteq \mathcal{T}$: That is, for any other topology $\mathcal{T}'$ on X with $\mathcal{B} \subseteq \mathcal{T}'$ we have that $\mathcal{T} \subseteq \mathcal{T}'$. The topology $\mathcal{T}$ can be defined simply as the intersection of all topologies on X that contain $\mathcal{B}$ (here it is vital that the intersection of a collection of topologies is again a topology). The set $\mathcal{B}$ becomes a subbasis for the topology $\mathcal{T}$, i.e. the collection of all finite intersections of sets from $\mathcal{B}$, together with X itself, forms a basis for $\mathcal{T}$.

Definition 1.1.6. Let X be a set, let $(Y_i)_{i \in I}$ be a collection of topological spaces and let $(f_i)_{i \in I}$ be a collection of functions $f_i \colon X \to Y_i$. The weakest topology on X that makes every function f_i $(i \in I)$ continuous is called the *initial topology* on X with respect to $(f_i)_{i \in I}$.

Remark. In the setting of Definition 1.1.6, consider the collection

$$\mathcal{B} = \{f_i^{-1}(U) : i \in I \text{ and } U \text{ is open in } Y_i\}. \tag{1.1}$$

Then the initial topology on X is the weakest topology that contains $\mathcal{B}$, so $\mathcal{B}$ is a subbasis for this topology. It follows that a *basis* for the topology is given by finite intersections of elements from $\mathcal{B}$, and that any open set can be written as a union of sets from this basis.

Example 1.1.7. (a) Let X be a topological space and let S be a subset of X. Then the initial topology on S with respect to the inclusion map $i\colon S \hookrightarrow X$ is the subspace topology: It is the weakest topology on S that contains the sets $i^{-1}(U) = U \cap S$ for every open set U in X. In this case this collection is already a topology.

(b) Let $(X_i)_{i\in I}$ be a collection of topological spaces and set $X = \prod_{i\in I} X_i$. Then the weakest topology on X such that all the projections $X \to X_i$, $(x_j)_{j\in I} \mapsto x_i$ are continuous is the product topology on X.

Proposition 1.1.8. *Let X be equipped with the initial topology with respect to a collection $(f_i)_{i\in I}$ of maps $f_i\colon X \to Y_i$. Suppose we are given for each $i \in I$ and $y \in Y_i$ a neighborhood basis $\mathcal{B}_y$ for y. Then a neighborhood basis at $x \in X$ is given by*

$$\mathcal{B}_x = \left\{ \bigcap_{i\in I_0} f_i^{-1}(B_i) : \ I_0 \subseteq I \textit{ is finite and } B_i \in \mathcal{B}_{f_i(x)} \textit{ for each } i \in I_0 \right\}.$$

Proof. Let $N \in \mathcal{N}_x$. Then there exists an open set $U \subseteq N$ with $x \in U$. Since the collection $\mathcal{B}$ from (1.1) is a subbasis for the topology, we can find finitely many elements from $\mathcal{B}$ that all contain x such that their intersection is contained in U. Thus we have a finite set $I_0 \subseteq I$ and open sets $U_i \subseteq Y_i$ for each $i \in I_0$ such that $x \in \bigcap_{i\in I_0} f_i^{-1}(U_i) \subseteq U$. By definition of neighborhood basis we can for each $i \in I_0$ find $B_i \in \mathcal{B}_{f_i(x)}$ such that $f_i(x) \in B_i \subseteq U_i$. But then $x \in \bigcap_{i\in I_0} f_i^{-1}(B_i) \subseteq U$. This finishes the proof. ∎

Note that in particular, we can choose $\mathcal{B}_y = \mathcal{N}_y$ for every $y \in Y_i$ in Proposition 1.1.8.

Example 1.1.9. Let $(X_i)_{i\in I}$ be a collection of topological spaces. Set $X = \prod_{i\in I} X_i$ and let $x = (x_i)_{i\in I} \in X$. Denote by $p_i\colon X \to X_i$ the projection onto X_i ($i \in I$). By Proposition 1.1.8 we can describe a neighborhood basis at x as follows: As I_0 ranges over finite subsets of I and N_i is a neighborhood of x_i for each $i \in I_0$, the sets

$$\bigcap_{i\in I_0} p_i^{-1}(N_i) = \prod_{i\in I} B_i \quad \text{where } B_i = N_i \text{ if } i \in I_0 \text{ and } B_i = X_i \text{ otherwise}$$

form a neighborhood basis at x.

1.2 Nets

Recall that a preorder on a set X is a relation $\leq$ on X which is reflexive ($x \leq x$ for all $x \in X$) and transitive ($x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in X$).

Definition 1.2.1. A *directed set* is a nonempty set Λ together with a preorder $\leq$ on X such that the following holds: For every $x, y \in \Lambda$ there exists $z \in \Lambda$ such that $x \leq z$ and $y \leq z$.

Example 1.2.2. (a) Every totally ordered set is a directed set. In particular $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are directed sets with their usual ordering.

(b) Let X be a topological space and fix a point $x \in X$. Let $\mathcal{B}_x$ be a neighborhood basis at x. Then $\mathcal{B}_x$ is a directed set with respect to *reverse inclusion* by Proposition 1.1.5 (a). This applies in particular to $\mathcal{N}_x$, the neighborhood filter at x.

(c) Let $[a, b]$ be an interval. Recall that a tagged partition $P = ((x_i)_{i=0}^{n}, (t_i)_{i=0}^{n-1})$ of $[a, b]$ is a finite sequence of the form $a = x_0 < x_1 < \cdots < x_n = b$ together with points $t_i \in [x_i, x_{i+1}]$ for each $0 \leq i \leq n-1$. The mesh of P is the maximum of the numbers $x_{i+1} - x_i$ for $0 \leq i \leq n-1$. Let $\mathcal{P}$ be the set of all tagged partitions of $[a, b]$. We introduce an ordering on Λ as follows: $((x_i)_{i=0}^{m}, (t_i)_{i=0}^{m-1}) \leq ((y_i)_{i=0}^{n}, (s_i)_{i=0}^{n-1})$ if and only if for every $0 \leq i \leq n$ there exists an integer $r(i)$ such that $x_i = y_{r(i)}$ and $t_i = s_j$ for some $r(i) \leq j \leq r(i+1)$. One can check that Λ is a directed set with respect to this ordering. This comes from the fact that every pair of tagged partitions have a *common refinement*.

Definition 1.2.3. Let X be a topological space. A *net* in X is a collection of elements $(x_\lambda)_{\lambda \in \Lambda}$ in X indexed by some directed set Λ.

Let $(x_\lambda)_{\lambda \in \Lambda}$ be a net in X and let $S \subseteq X$. We say that $(x_\lambda)_{\lambda \in \Lambda}$ is *frequently* in S if for every $\lambda_0 \in \Lambda$ there exists $\lambda \geq \lambda_0$ such that $x_\lambda \in S$. We say that $(x_\lambda)_{\lambda \in \Lambda}$ is *eventually* in S if there exists $\lambda_0 \in \Lambda$ such that for all $\lambda \geq \lambda_0$ we have $x_\lambda \in S$. We say that $(x_\lambda)_{\lambda \in \Lambda}$ *converges to* x if $(x_\lambda)_{\lambda \in \Lambda}$ is eventually in N for every $N \in \mathcal{N}_x$. We call x an *cluster point* for $(x_\lambda)_{\lambda \in \Lambda}$ if $(x_\lambda)_{\lambda \in \Lambda}$ is frequently in N for every $N \in \mathcal{N}_x$.

Remark. Note that in the definition of convergence and cluster point we can interchange $\mathcal{N}_x$ with any other neighborhood basis $\mathcal{B}_x$ at x and get an equivalent definition. In particular it can be convenient to use the neighborhood basis $\mathcal{U}_x$ of open neighborhoods of x.

Example 1.2.4. (a) A net in X indexed by the directed set $\mathbb{N}$ is the same as a sequence in X, and the definition of convergence in this case is exactly the same as the definition of convergence of a sequence.

(b) Let x be a point in a topological space X and let $\mathcal{B}_x$ be a neighborhood basis at x. Consider $\Lambda = \mathcal{B}_x$ as a directed set with respect to reverse inclusion. Let $(x_B)_{B \in \Lambda}$ be a net in X with the property that $x_B \in B$ for every $B \in \Lambda$. Then the net $(x_B)_B$ converges to x and the proof is almost tautological: For every neighborhood N of x we simply take some $B_0 \in \mathcal{B}_x$ with $B_0 \subseteq N$. Then whenever $B \subseteq B_0$ (remember that we are using reverse inclusion as the order relation on Λ) we have that $x_B \in B \subseteq B_0 \subseteq N$.

(c) Let Λ be the directed set from Example 1.2.2 (c). Given a bounded function $f \colon [a, b] \to \mathbb{R}$ we construct a net of real numbers indexed by Λ as follows: Given $P = ((x_i)_{i=0}^{n}, (t_i)_{i=0}^{n-1}) \in \Lambda$ set

$$I_P = \sum_{i=0}^{n-1} f(t_i)(x_{i+1} - x_i).$$

If the net $(I_P)_{P \in \Lambda}$ converges, we say that f is *Riemann integrable*, and denote the limit of $(I_P)_P$ by $\int_a^b f(x)\, \mathrm{d}x$.

Proposition 1.2.5. *Let X and Y be topological spaces, let $f\colon X \to Y$ be a function and let $x \in X$. Then the following are equivalent:*

(a) f is continuous at x.

(b) For every net $(x_\lambda)_{\lambda\in\Lambda}$ in X converging to x, the net $(f(x_\lambda))_{\lambda\in\Lambda}$ converges to $f(x)$.

Proof. $(a) \Rightarrow (b)$: Suppose that f is continuous at x. Let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in X converging to $x \in X$. Let N be a neighborhood of $f(x)$. Then $f^{-1}(N)$ is a neighborhood of x, so by assumption there exists λ_0 with

$$\lambda \geq \lambda_0 \implies x_\lambda \in f^{-1}(N).$$

But if $x_\lambda \in f^{-1}(N)$ then $f(x_\lambda) \in N$. Hence the net $(f(x_\lambda))_{\lambda\in\Lambda}$ converges to $f(x)$.

$(b) \Rightarrow (a)$: We do a contrapositive proof. Suppose f is not continuous at x. Then we can find some neighborhood N' of $f(x)$ such that no neighborhood N of x has the property that $f(N) \subseteq N'$. In other words, for every $N \in \mathcal{N}_x$ there exists $x_N \in N$ with $f(x_N) \notin N'$. Consider $\Lambda = \mathcal{N}_x$ as a directed set with respect to reverse inclusion and consider the net $(x_N)_{N\in\Lambda}$. As we saw in Example 1.2.4 (b), $(x_N)_N$ converges to x. However, $(f(x_N))_{N\in\Lambda}$ cannot converge to $f(x)$ since the neighborhood N' of $f(x)$ has the property that $f(x_N) \notin N'$ for all $N \in \Lambda$. This finishes the contrapositive proof. ■

Proposition 1.2.6. *Let X be a topological space. Then the following are equivalent:*

(a) X is Hausdorff.

(b) Limits of nets are unique: That is, whenever $(x_\lambda)_{\lambda\in\Lambda}$ is a net and both x and y are limits of $(x_\lambda)_{\lambda\in\Lambda}$, then $x = y$.

Proof. Left as an exercise. ■

If X is Hausdorff and $(x_\lambda)_{\lambda\in\Lambda}$ is a convergent net, we can thus speak of *the* limit of $(x_\lambda)_\lambda$ and write $x = \lim_{\lambda\in\Lambda} x_\lambda$.

Proposition 1.2.7. *Let X be a topological space, let S be a subset of X and let $x \in X$. Then the following are equivalent:*

(a) x is in $\operatorname{Cl} S$, the closure of S.

(b) There exists a net $(x_\lambda)_{\lambda\in\Lambda}$ in S that converges to x.

Proof. $(a) \Rightarrow (b)$: Suppose $x \in \operatorname{Cl} S$. Let $\Lambda = \mathcal{N}_x$ with respect to reverse inclusion. Then for every $N \in \mathcal{N}_x$ we have that $N \cap S \neq \emptyset$ so we can pick some $x_N \in N \cap S$. By Example 1.2.4 the net $(x_N)_{N\in\Lambda}$ converges to x.

$(b) \Rightarrow (a)$: Suppose that $(x_\lambda)_{\lambda\in\Lambda}$ is a net in S that converges to x. For a contradiction, suppose $x \notin \operatorname{Cl} S$. Then there exists a closed subset $C \supseteq S$ such that $x \notin C$. But then C^c is an open neighborhood of x, so there exists $\lambda_0 \in \Lambda$ for which $x_\lambda \in C^c$ whenever $\lambda \geq \lambda_0$. But then we have both $x_{\lambda_0} \in S$ and $x_{\lambda_0} \in C^c \subseteq S^c$ which is a contradiction. We conclude that $x \in \operatorname{Cl} S$. ■

Proposition 1.2.8. *Let X be equipped with the initial topology with respect to a family of functions $f_i\colon X \to Y_i$ $(i \in I)$. Let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in X and let $x \in X$. Then the following are equivalent:*

(a) $(x_\lambda)_{\lambda\in\Lambda}$ *converges to x.*

(b) For every $i \in I$ the net $(f_i(x_\lambda))_{\lambda\in\Lambda}$ converges to $f_i(x)$.

Proof. Exercise. ∎

1.3 Subnets and compactness

Recall that a sequence $(y_n)_{n\in\mathbb{N}}$ is said to be a subsequence of a sequence $(x_n)_{n\in\mathbb{N}}$ if there exists a monotone injective function $\iota\colon \mathbb{N} \to \mathbb{N}$ such that $y_n = x_{\iota(n)}$ for all $n \in \mathbb{N}$. Usually one writes $n_k = \iota(k)$ for $k \in \mathbb{N}$ so that $(y_n)_n$ becomes $(x_{n_k})_{k\in\mathbb{N}}$. The right generalization to nets turns out to be somewhat more complicated.

Definition 1.3.1. Let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in a topological space X. A *subnet* of $(x_\lambda)_{\lambda\in\Lambda}$ is a net $(y_\gamma)_{\gamma\in\Gamma}$ together with a function $\iota\colon \Gamma \to \Lambda$ such that the following properties are satisfied:

(a) ι is *monotone*, i.e. $\gamma \le \gamma'$ implies $\iota(\gamma) \le \iota(\gamma')$ for all $\gamma, \gamma' \in \Gamma$.

(b) ι is *cofinal*, i.e. for every $\lambda \in \Lambda$ there exists $\gamma \in \Gamma$ such that $\iota(\gamma) \ge \lambda$.

(c) $y_\gamma = x_{\iota(\gamma)}$ for all $\gamma \in \Gamma$.

We often use the more compact notation $(x_{\iota(\gamma)})_{\gamma\in\Gamma}$ for a subnet of $(x_\lambda)_{\lambda\in\Lambda}$.

Example 1.3.2. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a topological space X. Consider $\mathbb{R}$ as a directed set with its usual ordering. Define a map $\iota\colon \mathbb{R}^+ \to \mathbb{N}$ by $\iota(x) = \lfloor x \rfloor$, the floor function. Then ι satisfies *(a)* and *(b)* in Definition 1.3.1. Set $y_r = x_{\iota(r)}$ for every $r \in \mathbb{R}^+$. Then $(y_r)_{r\in\mathbb{R}^+}$ becomes a subnet of $(x_n)_{n\in\mathbb{N}}$, but it is not a subsequence of $(x_n)_{n\in\mathbb{N}}$.

Proposition 1.3.3. *Let X be a topological space and let $x \in X$. If $(x_\lambda)_{\lambda\in\Lambda}$ is a net in X that converges to x, then every subnet of $(x_\lambda)_{\lambda\in\Lambda}$ also converges to x.*

Proof. Let $(y_\gamma)_{\gamma\in\Gamma}$ be a subnet of $(x_\lambda)_{\lambda\in\Lambda}$ with associated function $\iota\colon \Gamma \to \Lambda$ and let N be a neighborhood of x. Since $(x_\lambda)_{\lambda\in\Lambda}$ converges to x, we can find $\lambda_0 \in \Lambda$ such that $x_\lambda \in N$ whenever $\lambda \ge \lambda_0$. By cofinality of ι we can find $\gamma_0 \in \Gamma$ such that $\iota(\gamma_0) \ge \lambda_0$. Now if $\gamma \ge \gamma_0$ then $\iota(\gamma) \ge \iota(\gamma_0) \ge \lambda_0$ by monotonicity. Thus $\gamma \ge \gamma_0$ implies that $y_\gamma = x_{\iota(\gamma)} \in N$. Since N was arbitrary, we conclude that $(y_\gamma)_{\gamma\in\Gamma}$ converges to x. ∎

Proposition 1.3.4. *Let $\mathcal{B}$ be a nonempty set of subsets of X which is a directed set with respect to reverse inclusion. If $(x_\lambda)_{\lambda\in\Lambda}$ is a net that is frequently in every $B \in \mathcal{B}$, then there exists a subnet of $(x_\lambda)_{\lambda\in\Lambda}$ that is eventually in every $B \in \mathcal{B}$.*

Proof. Consider the set

$$\Gamma = \{(\lambda, B) \in \Lambda \times \mathcal{B} : x_\lambda \in B\} \subseteq \Lambda \times \mathcal{B}$$

equipped with the order where we declare $(\lambda, B) \leq (\lambda', B')$ iff both $\lambda \leq \lambda'$ and $B \supseteq B'$.

We claim that Γ is a directed set: Suppose $(\lambda, B), (\lambda', B') \in \Gamma$. Let $B'' \in \mathcal{B}$ be such that $B'' \subseteq B \cap B'$. Since $(x_\lambda)_{\lambda\in\Lambda}$ is frequently in B'' by assumption we can find $\lambda'' \in \Lambda$ with $\lambda'' \geq \lambda$, $\lambda'' \geq \lambda'$ such that $x_{\lambda''} \in B''$. Thus $(\lambda'', B'') \geq (\lambda, B)$ and $(\lambda'', B'') \geq (\lambda', B')$ which shows that Γ is a directed set.

Define a map $\iota \colon \Gamma \to \Lambda$ by $\iota(\lambda, B) = \lambda$. Then ι is obviously monotone. It is also cofinal: If $\lambda \in \Lambda$, then pick some $B \in \mathcal{B}$. By the assumption on $(x_\lambda)_{\lambda\in\Lambda}$ there exists $\lambda' \in \Lambda$ such that $\lambda' \geq \lambda$ and $x_{\lambda'} \in B$. Hence $(\lambda', B) \in \Gamma$ and $\iota(\lambda', B) = \lambda' \geq \lambda$. Thus, we have proved that $(x_{\iota(\lambda,B)})_{(\lambda,B)\in\Gamma}$ is a subnet of $(x_\lambda)_{\lambda\in\Lambda}$. Furthermore, if $B_0 \in \mathcal{B}$ then pick any $\lambda_0 \in \Lambda$ such that $x_{\lambda_0} \in B_0$. Then $(\lambda_0, B_0) \in \Gamma$, and if $(\lambda, B) \in \Gamma$ with $(\lambda, B) \geq (\lambda_0, B_0)$ then $\lambda \geq \lambda_0$ and $x_{\iota(\lambda,B)} = x_\lambda \in B \subseteq B_0$. This shows that the net $(x_{\iota(\lambda,B)})_{(\lambda,B)\in\Gamma}$ is eventually in B_0. ■

Corollary 1.3.5. *Let X be a topological space, let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in X and let $x \in X$. The following are equivalent:*

(a) x is an cluster point for $(x_\lambda)_{\lambda\in\Lambda}$.

(b) $(x_\lambda)_{\lambda\in\Lambda}$ has a subnet that converges to x.

Proof. (a) ⇒ (b): Use Proposition 1.3.4 on $\mathcal{B} = \mathcal{N}_x$.

(b) ⇒ (a): Let $(x_{\iota(\gamma)})_{\gamma\in\Gamma}$ be a subnet that converges to x. Let $N \in \mathcal{N}_x$ and $\lambda_0 \in \Lambda$. By cofinality we can find $\gamma_0 \in \Gamma$ such that $\iota(\gamma_0) \geq \lambda_0$, and by convergence we can find $\gamma_1 \in \Gamma$ such that $x_{\iota(\gamma)} \in N$ whenever $\gamma \geq \gamma_1$. Let $\gamma \in \Gamma$ satisfy $\gamma \geq \gamma_0$ and $\gamma \geq \gamma_1$ and set $\lambda = \iota(\gamma)$. Then $\lambda \geq \lambda_0$ and $x_\lambda \in N$, which shows that x is a cluster point for $(x_\lambda)_{\lambda\in\Lambda}$. ■

Recall that a topological space X is *compact* if every open cover of X admits a finite subcover: That is, whenever $X = \bigcup_{i\in I} U_i$ for a collection $(U_i)_{i\in I}$ of open sets in X, then there exists a finite subset $I_0 \subseteq I$ such that $X = \bigcup_{i\in I_0} U_i$. The equivalent contrapositive formulation of this definition is called the *finite intersection property*: Whenever $(C_i)_{i\in I}$ is a collection of closed subsets of X with the property that $\bigcap_{i\in I_0} C_i \neq \emptyset$ for all finite subsets I_0 of I, then $\bigcap_{i\in I} C_i \neq \emptyset$.

Lemma 1.3.6. *If $(x_\lambda)_{\lambda\in\Lambda}$ is a net in a topological space X, then its set of cluster points coincides with*

$$\bigcap_{\lambda\in\Lambda} \mathrm{Cl}\{x_{\lambda'} : \lambda' \geq \lambda\}.$$

Proof. Note that x is a cluster point of $(x_\lambda)_\lambda$ if and only if the following holds for all $\lambda \in \Lambda$: Every neighborhood N of x intersects the set $\{x_{\lambda'} : \lambda' \geq \lambda\}$. This is equivalent to x being in the closure of $\{x_{\lambda'} : \lambda' \geq \lambda\}$ for every $\lambda \in \Lambda$, which proves the lemma. ■

Proposition 1.3.7. *Let X be a topological space. The following are equivalent:*

(a) X is compact.

(b) Every net in X has a cluster point.

(c) *Every net in X has a convergent subnet.*

Proof. (a) $\Rightarrow$ (b): Suppose X is compact and let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in X. For each $\lambda \in \Lambda$ set $C_\lambda = \mathrm{Cl}\{x_{\lambda'} : \lambda' \geq \lambda\}$. Then $(C_\lambda)_{\lambda\in\Lambda}$ is a collection of nonempty closed sets with the finite intersection property: Indeed, if $\Lambda_0 \subseteq \Lambda$ is finite then we can find $\lambda' \in \Lambda$ such that $\lambda' \geq \lambda$ for every $\lambda \in \Lambda_0$, in which case

$$\emptyset \neq C_{\lambda'} \subseteq \bigcap_{\lambda\in\Lambda_0} C_\lambda.$$

Since X is compact we have that the intersection $\bigcap_{\lambda\in\Lambda} C_\lambda$ is nonempty. By Lemma 1.3.6 this implies that $(x_\lambda)_\lambda$ has a cluster point.

(b) $\Leftrightarrow$ (c): By Corollary 1.3.5, a given net has a cluster point if and only if it has a convergent subnet. Thus, the statement that all nets have cluster points is equivalent to the statement that all nets have convergent subnets.

(b) $\Rightarrow$ (a): We prove the contrapositive statement. Let $(U_i)_{i\in I}$ be an open cover of X and assume for a contradiction that it does not have a finite subcover. Let Λ be the set of all finite subsets of I ordered by inclusion. Then Λ is a directed set. If $J \in \Lambda$ then by assumption we can pick some $x_J \in X \setminus \bigcup_{i\in J} U_i$. Thus we have a net $(x_J)_{J\in\Lambda}$. Let $x \in X$ be arbitrary. We claim that x cannot be a cluster point for $(x_J)_{J\in\Lambda}$: Indeed, since $(U_i)_{i\in I}$ covers X there exists some $i \in I$ such that $x \in U_i$. Thus U_i is a neighborhood of x, yet when $J \supseteq \{i\}$ we have that $x_J \notin U_i$. ■

We have now come to our first theorem, namely Tychonoff's theorem:

Theorem 1.3.8 (Tychonoff's Theorem). *Let $(X_i)_{i\in I}$ be a collection of compact topological spaces. Then the product space $\prod_{i\in I} X_i$ is compact (when equipped with the product topology).*

To prove Tychonoff's theorem we will need the following lemma:

Lemma 1.3.9. *Let X and Y be topological spaces, and let $(z_\lambda)_{\lambda\in\Lambda}$ be a net in the product space $X \times Y$. Denote by $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ the coordinate projections. If both of the nets $(p_X(z_\lambda))_{\lambda\in\Lambda}$ and $(p_Y(z_\lambda))_{\lambda\in\Lambda}$ have cluster points, then $(z_\lambda)_{\lambda\in\Lambda}$ has a cluster point.*

Proof. Let $x \in X$ (resp. $y \in Y$) be a cluster point for $(p_X(z_\lambda))_{\lambda\in\Lambda}$ (resp. $(p_Y(z_\lambda))_{\lambda\in\Lambda}$). By Proposition 1.3.7, there exists a subsequence $(z_{\iota(\gamma)})_{\gamma\in\Gamma}$ of $(z_\lambda)_\lambda$ such that $(p_X(z_{\iota(\gamma)}))_\gamma$ converges to x. For the same reason there exists a subsequence $(z_{\kappa(\iota(\alpha))})_{\alpha\in A}$ of $(z_\lambda)_\lambda$ such that $(p_Y(z_{\kappa(\iota(\alpha))})_\alpha$ converges to y. Since $(p_X(z_{\kappa(\iota(\alpha))})_\alpha$ is a subsequence of $(p_X(z_{\iota(\gamma)}))_\gamma$, it must also converge to x by Proposition 1.3.3.

We now claim that $(z_{\kappa(\iota(\alpha))})_\alpha$ converges to (x, y): Indeed, let N be a neighborhood of (x, y). By definition of the product topology, we can assume that $N = N_1 \times N_2$ where N_1 is a neighborhood of x and N_2 is a neighborhood of y. Thus, we can find $\alpha_1, \alpha_2 \in A$ such that $p_X(z_{\kappa(\iota(\alpha))}) \in N_1$ when $\alpha \geq \alpha_1$ and $p_Y(z_{\kappa(\iota(\alpha))}) \in N_2$ when $\alpha \geq \alpha_2$. Picking $\alpha_0 \in A$ with $\alpha_0 \geq \alpha_1$ and $\alpha_0 \geq \alpha_2$, we have that $z_{\kappa(\iota(\alpha))} \in N_1 \times N_2$ when $\alpha \geq \alpha_0$. Thus $(z_{\kappa(\iota(\alpha))})_\alpha$ converges to (x, y). We have shown that $(z_\lambda)_\lambda$ has a convergent subsequence, and therefore a cluster point by Proposition 1.3.7. ■

Proof of Theorem 1.3.8. Let $(z_\lambda)_{\lambda\in\Lambda}$ be a net in $X := \prod_{i\in I} X_i$. We will show that $(z_\lambda)_{\lambda\in\Lambda}$ has a cluster point. It will then follow from Proposition 1.3.7 that X is compact.

Given subsets J and K of I with $J \supseteq K$, we denote by $p_{J,K}$ the projection from $\prod_{i\in J} X_i$ to $\prod_{i\in K} X_i$. We consider the set $\mathcal{M}$ of pairs (J,x) where J is a subset of I and $x \in \prod_{i\in J} X_i$ is a cluster point of the net $(p_{I,J}(z_\lambda))_{\lambda\in\Lambda}$. Define a relation $\leq$ on $\mathcal{M}$ via

$$(J,x) \leq (K,y) \quad \text{if and only if} \quad J \subseteq K \text{ and } p_{K,J}(y) = x.$$

We leave it as an exercise to check that this defines a partial order on $\mathcal{M}$. Our strategy is now to employ Zorn's lemma. We must therefore show that every chain in $\mathcal{M}$ has an upper bound.

Let $\mathcal{C}$ be a chain in $\mathcal{M}$, that is, a totally ordered subset. Set $K = \bigcup_{(J,x)\in\mathcal{C}} J$. Define an element $y = (y_i)_{i\in K} \in \prod_{i\in K} X_i$ as follows: For $i \in K$ we have that there exists some $(J,x) \in \mathcal{C}$ such that $i \in J$. Set $y_i = x_i$. We show that this is well-defined: Suppose $(J',x') \in \mathcal{C}$ is also such that $i \in J'$. Since $\mathcal{C}$ is a chain we have either $(J,x) \leq (J',x')$ or $(J',x') \leq (J,x)$. If the former is the case then $J \subseteq J'$ and $p_{J',J}(x') = x$ which implies that $x_i = x'_i$ since $i \in J \cap J'$. A similar argument holds for the case $(J',x') \leq (J,x)$. Hence (J,y) is well-defined.

We must now show that (K,y) is an element of $\mathcal{M}$, i.e. that y is a cluster point of $(p_{I,K}(z_\lambda))_{\lambda\in\Lambda}$. Let $N \in \mathcal{N}_y$ and let $\lambda_0 \in \Lambda$. By Proposition 1.1.8 we can assume without loss of generality that $N = \prod_{i\in K} N_i$ where $N_i \in \mathcal{N}_{y_i}$ for each $i \in K$ and there exists a finite subset K_0 of K such that $N_i = X_i$ when $i \notin K_0$. Now every $i \in K_0$ belongs to some $J_i \subseteq K$ such that $(J_i, y_i) \in \mathcal{C}$ for some $y_i \in \prod_{j\in J_i} X_j$. Since $\mathcal{C}$ is a chain and K_0 is finite, we can pick a maximum among these (J_i, y_i) as i runs through K_0. Call it (K', y'). Then $K_0 \subseteq K'$. Now y' is a cluster point of $(p_{I,K'}(z_\lambda))_{\lambda\in\Lambda}$ by definition. Thus $p_{K',K_0}(y')$ is a cluster point of the net $(p_{I,K_0}(z_\lambda))_\lambda = (p_{K',K_0}(p_{I,K'}(z_\lambda)))_\lambda$. Consequently we can find $\lambda \geq \lambda_0$ such that $p_{I,K_0}(z_\lambda) \in \prod_{i\in K_0} N_i$. But then $p_{I,K}(z_\lambda) \in N$ since $N_i = X_i$ for all $i \notin K_0$. Hence y is a cluster point of $(p_{I,K}(z_\lambda))_{\lambda\in\Lambda}$.

By Zorn's lemma, we conclude that $\mathcal{M}$ has a maximal element, say (I',x). We want to show that $I' = I$, from which it will follows that x is a cluster point for $(z_\lambda)_{\lambda\in\Lambda} = (p_{I,I}(z_\lambda))_{\lambda\in\Lambda}$. Assume for a contradiction that there exists $j \in I \setminus I'$. Then by compactness of X_j the net $(p_{I,\{j\}}(z_\lambda))_\lambda$ has a cluster point. Since $(p_{I,I'}(z_\lambda))_\lambda$ also has a cluster point, it follows from Lemma 1.3.9 that the net $(p_{I,I'\cup\{j\}}(z_\lambda))_\lambda$ has a cluster point. This contradicts the maximality of (I',x). Thus $I = I'$, and the proof is finished. ∎

Chapter 2

Topological vector spaces

2.1 Convex sets and semi-norms

Throughout this chapter, $\mathbb{F}$ will denote either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. All vector spaces will be over $\mathbb{F}$ unless stated otherwise.

Definition 2.1.1. Let X be a vector space and let $x, y \in X$. The *line segment between x and y* is the set

$$[x, y] = \{(1-t)x + ty : 0 \leq t \leq 1\} \subseteq X.$$

A subset C of X is called *convex* if for every pair $x, y \in C$, the line segment between x and y is contained in C.

By induction one verifies that a convex set C contains every element of the form

$$\sum_{i=1}^{n} \lambda_i x_i$$

where $x_1, \ldots, x_n \in C$ and $\lambda_1, \cdots \lambda_n$ are nonnegative numbers such that $\sum_{i=1}^{n} \lambda_i = 1$. Such a sum $\sum_i \lambda_i x_i$ is called a *convex* combination of $x_1, \ldots, x_n$.

It can be readily verified that the intersection of a collection of convex sets is convex. Given a subset S of X, we define the *convex hull* of S to be the intersection of all convex sets containing S, i.e. the smallest convex set in X that contains S. We denote the convex hull of S by $\operatorname{co} S$.

Proposition 2.1.2. *Let X be a vector space and let S be a subset of X. Then the convex hull of S consists precisely of all convex combinations of elements from S.*

Proof. Denote by C the convex hull of S and by C' the set of all convex combinations of elements from S. Then $S \subseteq C'$ (for any $x \in S$ pick $n = 1$, $x_1 = x$ and $\lambda_1 = 1$). We will show that C' is convex: Let $x, y \in C'$ and $t \in [0, 1]$. Then $x = \sum_{i=1}^{m} \lambda_i x_i$ and $y = \sum_{j=1}^{n} \mu_j y_j$ for some $x_i, y_j \in S$ and $\lambda_i, \mu_j \geq 0$ with $\sum_i \lambda_i = \sum_j \mu_j = 1$. Since

$$\sum_{i=1}^{m} (1-t)\lambda_i + \sum_{j=1}^{n} t\mu_j = t \cdot 1 + (1-t) \cdot 1 = 1$$

we have that $(1-t)x+ty = \sum_i (1-t)\lambda_i x_i + \sum_j t\mu_j y_j$ is a convex combination of the elements $x_i, y_j \in S$, hence $(1-t)x+ty \in C'$.

Since C' is convex and contains S this shows that $C \subseteq C'$. On the other hand, every convex set that contains S must contain all convex combinations of elements from S, so $C' \subseteq C$. ■

Definition 2.1.3. Let X be a vector space. A subset S of X is called

(a) *balanced* if for all $x \in S$ and $\lambda \in \mathbb{F}$ with $|\lambda| \leq 1$ we have that $\lambda x \in S$,

(b) *absorbing* if for all $x \in X$ there exists $\epsilon > 0$ such that $\lambda x \in S$ for all $\lambda \in \mathbb{F}$ with $|\lambda| < \epsilon$.

Note that absorbing sets contain the zero vector. The same holds for nonempty balanced sets. Furthermore, any intersection of balanced sets is balanced. However, only *finite* intersections of absorbing sets are absorbing in general. On the other hand, any union of absorbing sets is absorbing.

For subsets S and T of a vector space X and $\lambda \in \mathbb{F}$ we define

$$S + T = \{x + y : x \in S, y \in T\},$$
$$\lambda S = \{\lambda x : x \in S\}.$$

In particular $-S = \{-x : x \in S\}$ and $S - T = S + (-T)$. We also set $x + S = \{x\} + S$ for $x \in X$ and call it the *translate* of S by x.

Remark. It is *not* true in general that e.g. $S + S = 2S$. The former consists of all possible sums $x + y$ for $x, y \in S$ while the latter is just all multiples $2x$ where $x \in S$. However, if K is a convex set and $\lambda, \mu \geq 0$ then $(\lambda + \mu)K = \lambda K + \mu K$.

Proposition 2.1.4. *Let S and T be subsets of a vector space X and let $\lambda \in \mathbb{F}$. Then the following hold:*

(a) If S and T are convex then $S + T$ and λS are convex.

(b) If S is absorbing and $\lambda \neq 0$ then λS is absorbing.

(c) If S is balanced then λS is balanced.

Proof. Exercise. ■

A standard example of a set that is convex, balanced and absorbing is the open or closed unit ball in a normed space. More generally, we shall see that semi-norms give natural examples of such sets.

Definition 2.1.5. Let X be a vector space. A *semi-norm* on X is a function $\sigma\colon X \to \mathbb{R}$ that satisfies the following properties:

(a) For every $x, y \in X$ we have $\sigma(x + y) \leq \sigma(x) + \sigma(y)$.

(b) For every $x \in X$ and $\lambda \in \mathbb{F}$ we have that $\sigma(\lambda x) = |\lambda|\sigma(x)$.

Example 2.1.6. (a) A norm σ on a vector space X is a semi-norm with the extra property that $\sigma(x) = 0$ implies $x = 0$ for all $x \in X$.

(b) Let X be a vector space and let $\phi\colon X \to \mathbb{F}$ be a linear functional. Then $|\phi|\colon X \to [0,\infty)$ given by $|\phi|(x) = |\phi(x)|$ defines a semi-norm on X.

Remark. Note that $\sigma(0) = \sigma(0x) = |0|\sigma(x) = 0$ and $0 = \sigma(x-x) \leq \sigma(x) + |-1|\sigma(x) = 2\sigma(x)$, so semi-norms are nonnegative. If σ satisfies (a) above but instead of (b) only the weaker property $\sigma(\lambda x) = \lambda\sigma(x)$ for all $\lambda \geq 0$, then σ is called a *sublinear functional.* In contrast to semi-norms, sublinear functionals may take negative values. Also, if p is a sublinear functional then $\sigma(x) = \max\{p(x), p(-x)\}$ is a semi-norm.

Just like norms, a semi-norm σ on a vector space satisfies the reverse triangle inequality:

$$|\sigma(x) - \sigma(y)| \leq \sigma(x-y) \quad \text{for all } x, y \in X. \tag{2.1}$$

Given a semi-norm σ on X, $x_0 \in X$ and $r > 0$ we define the open and closed σ-ball of radius r and center x_0 respectively as

$$B_r^\sigma(x_0) = \{x \in X : \sigma(x - x_0) < r\},$$
$$\bar{B}_r^\sigma(x_0) = \{x \in X : \sigma(x - x_0) \leq r\}.$$

We also set $B_r^\sigma = B_r^\sigma(0)$ and $\bar{B}_r^\sigma = \bar{B}_r^\sigma(0)$. In particular we set $B^\sigma = B_1^\sigma$ and $\bar{B}^\sigma = \bar{B}_1^\sigma$. Note that $B_r^\sigma(x_0) = x_0 + rB^\sigma$ and $\bar{B}_r^\sigma(x_0) = x_0 + r\bar{B}^\sigma$.

Proposition 2.1.7. *Let σ be a semi-norm on a vector space X. Then the sets B^σ and $\bar{B}^\sigma$ are convex, balanced and absorbing.*

Proof. We show that B^σ is convex: Indeed, if $x, y \in B^\sigma$ and $t \in [0,1]$ then

$$\sigma((1-t)x + ty) \leq (1-t)\sigma(x) + t\sigma(y) < (1-t)\cdot 1 + t \cdot 1 = 1.$$

Moreover, B^σ is balanced: If $x \in B^\sigma$ and $|\lambda| \leq 1$ then $\sigma(\lambda x) = |\lambda|\sigma(x) < 1 \cdot 1 = 1$. Finally, B^σ is absorbing: If $x \in X$ and $\sigma(x) \neq 0$ then provided that $|\lambda| < \sigma(x)^{-1}$ we have $\sigma(\lambda x) < 1$, so $\lambda x \in B^\sigma$. The proofs for $\bar{B}^\sigma$ are analogous. ■

Definition 2.1.8. Let X be a vector space and let S be an absorbing subset X. The *gauge* associated to S is the function $m_S\colon X \to [0,\infty)$ via

$$m_S(x) = \inf\{\lambda > 0 : \lambda^{-1}x \in S\}.$$

Remark. Note that for absorbing sets S we have that for every $x \in X$ there exists some $r > 0$ such that $\lambda^{-1}x \in S$ when $|\lambda| > r$. Thus $m_S(x) \leq r$ so the gauge of m_S does in fact take values in $[0,\infty)$.

Another term for gauge used in the literature is Minkowski functional, although it is not a functional in the usual sense.

The following proposition describes some fundamental properties of gauges:

Proposition 2.1.9. *Let X be a vector space and let S be an absorbing subset of X. Then the following hold:*

(a) For every $x \in X$ and $\lambda \in [0,\infty)$ we have that

$$m_S(\lambda x) = \lambda m_S(x).$$

(b) If S is balanced then for every $x \in X$ and $\lambda \in \mathbb{F}$ we have that

$$m_S(\lambda x) = |\lambda| m_S(x).$$

(c) If S is convex then for every $x, y \in X$ we have that

$$m_S(x + y) \le m_S(x) + m_S(y).$$

(d) If S is convex then

$$B^{m_S} \subseteq S \subseteq \bar{B}^{m_S}.$$

Proof. (a): Let $x \in X$ and $\lambda \ge 0$ and note that

$$\{r > 0 : r^{-1}\lambda x \in S\} = \{\lambda r : r > 0, r^{-1}x \in S\} = \lambda\{r > 0 : r^{-1}x \in S\}.$$

Taking infima we obtain $m_S(\lambda x) = \lambda m_S(x)$.

(b): Let $x \in X$ and $\lambda \in \mathbb{F}$. Since S is balanced we have that $\mu x \in S$ if and only if $|\mu| x \in S$ for all $\mu \in \mathbb{F}$. Hence

$$\{r > 0 : r^{-1}\lambda x \in S\} = \{r > 0 : r^{-1}|\lambda| x \in S\} = |\lambda|\{r > 0 : r^{-1}x \in S\}.$$

Taking infima, the desired conclusion follows.

(c): Let $x, y \in X$. Suppose $r, s > 0$ with $r^{-1}x, s^{-1}y \in X$. Since S is convex we have that

$$\frac{1}{r+s}(x + y) = \frac{r}{r+s}(r^{-1}x) + \frac{s}{r+s}(s^{-1}y) \in S.$$

This shows that $m_S(x + y) \le r + s$. Since r and s were arbitrary we can take the infimum over all r and then over all s to obtain $m_S(x + y) \le m_S(x) + m_S(y)$.

(d): If $x \in S$ then since $1^{-1}x \in S$ we have that $m_S(x) \le 1$. Suppose that $m_S(x) < 1$. Then we can find some $\lambda \in \mathbb{F}$ with $0 < \lambda < 1$ such that $\lambda^{-1}x \in S$. But then $x = \lambda(\lambda^{-1}x) + (1-\lambda)0 \in S$. ■

Corollary 2.1.10. *Let X be a vector space and let C be a convex, absorbing subset of X. Then the gauge m_C is a sublinear functional on X. If C is balanced as well, then m_C is a semi-norm on X.*

Proposition 2.1.9 (d) shows that we cannot recover a general convex absorbing set completely from its gauge; that is, if C and D are convex absorbing sets with $m_C = m_D$ then we cannot necessarily conclude that $C = D$. Starting with a semi-norm however, we can completely recover it from its open or closed ball as the following proposition shows:

Proposition 2.1.11. *Let σ be a semi-norm on a vector space X. Let C be any convex, balanced and absorbing set such that*

$$B^\sigma \subseteq C \subseteq \bar{B}^\sigma.$$

Then $m_C = \sigma$, that is, σ equals the gauge associated to C.

Proof. Exercise. ■

2.2 Neighborhood bases at zero

Definition 2.2.1. Let X be a vector space over $\mathbb{F}$. A topology $\mathcal{T}$ on X is said to be *linear* if vector addition and scalar multiplication are continuous maps with respect to $\mathcal{T}$. That is, the maps

$$X \times X \to X, \quad (x, y) \mapsto x + y,$$
$$\mathbb{F} \times X \to X, \quad (\lambda, x) \mapsto \lambda x,$$

are continuous when X is equipped with the topology $\mathcal{T}$ and $X \times X$ and $\mathbb{F} \times X$ is equipped with the corresponding product topologies coming from $\mathcal{T}$ and the natural topology on $\mathbb{F}$. A vector space equipped with a linear topology is called a *topological vector space*.

Example 2.2.2. Let X be a normed vector space. Then X is a topological vector space when equipped with the topology induced by its norm. In particular, the finite-dimensional vector space $\mathbb{F}^n$ is a topological vector space with respect to the usual topology on $\mathbb{F}^n$, since this is the topology induced by any norm on $\mathbb{F}^n$.

Let X be a topological vector space. For each $x \in X$ define a map $h_x \colon X \to X$ by

$$h_x(y) = x + y.$$

Then by continuity of vector addition it follows that h_x is a continuous map. In fact, since h_x is invertible with $h_x^{-1} = h_{-x}$, each h_x is a homeomorphism. Since $x + S = \{x + y : y \in S\}$ is the image of S under the homeomorphism h_x, it follows that $x + U$ is open when U is open. Similar statements hold for closed sets, compact sets, and so on.

In particular, if $\mathcal{B}_0$ is a neighborhood basis for the zero vector and $x \in X$, then $\mathcal{B}_x \coloneqq \mathcal{B}_0 + x$ is a neighborhood basis for x. Thus a linear topology is uniquely determined by a neighborhood basis of the zero vector.

Proposition 2.2.3. *Let X be a topological vector space and let $\mathcal{N}_0$ be the neighborhood filter at 0. Then the following hold:*

(a) For every $N \in \mathcal{N}_0$ there exists $N' \in \mathcal{N}_0$ such that $N' + N' \subseteq N$.

(b) Every $N \in \mathcal{N}_0$ is absorbing.

(c) For every $N \in \mathcal{N}_0$ there exists a balanced $N' \in \mathcal{N}_0$ such that $N' \subseteq N$.

Proof. (*a*): Let $N \in \mathcal{N}_0$. Since vector multiplication is continuous at $(0, 0)$ and $0 + 0 = 0$, the inverse image of N under addition, i.e. $M = \{(x, y) \in X \times X : x + y \in N\}$ is a neighborhood of $(0, 0)$. By definition of product topology we can find $N_1, N_2 \in \mathcal{N}_0$ such that $N_1 \times N_2 \subseteq M$, i.e. $N_1 + N_2 \subseteq N$. Choosing $N' \in \mathcal{N}_0$ such that $N' \subseteq N_1$ and $N' \subseteq N_2$ we obtain $N' + N' \subseteq N$.

(*b*): Let $N \in \mathcal{N}_0$ and $x \in X$. Since scalar multiplication is continuous at $(0, x) \in \mathbb{F} \times X$ and $0x = 0$, it follows that $M = \{(\lambda, y) \in \mathbb{F} \times X : \lambda y \in N\}$ is a neighborhood of $(0, x)$. By definition of product topology we can find $r > 0$ and $N' \in \mathcal{N}_x$ such that $B_r(0) \times N' \subseteq M$. But then $\lambda x \in N$ whenever $|\lambda| < r$.

(*c*): Let $N \in \mathcal{N}_0$. By continuity of scalar multiplication at $(0, 0) \in \mathbb{F} \times X$ there exists $r > 0$ and $N' \in \mathcal{N}_0$ such that $B_r(0)N' \subseteq N$. Now $B_r(0)N'$ is a neighborhood of 0 (show this!) and if $\lambda \in \mathbb{F}$ and $|\lambda| \leq 1$ then $\lambda B_r(0) \subseteq B_r(0)$ so $\lambda B_r(0)N' \subseteq B_r(0)N'$. This shows that $B_r(0)N'$ is balanced. ∎

Corollary 2.2.4. *Let X be a topological vector space. Then X has a neighborhood basis at zero consisting of balanced, absorbing sets.*

Proof. Let $\mathcal{B}_0$ be the collection of balanced neighborhoods of 0. By Proposition 2.2.3 (b) every set in $\mathcal{B}_0$ is also absorbing. If $N \in \mathcal{N}_0$ then by Proposition 2.2.3 (c) there exists $B \in \mathcal{B}_0$ such that $B \subseteq N$. Thus $\mathcal{B}_0$ is a neighborhood basis at 0. ■

The following proposition gives conditions ensuring that a collection $\mathcal{B}_0$ of subsets of a vector space X is a neighborhood basis at zero for a (necessarily unique) linear topology on X.

Proposition 2.2.5. *Let X be a vector space and let $\mathcal{B}_0$ be a nonempty collection of subsets of X with the following properties:*

(a) Every $B \in \mathcal{B}_0$ is balanced and absorbing.

(b) For every $B, B' \in \mathcal{B}_0$ there exists $B'' \in \mathcal{B}_0$ such that $B'' \subseteq B \cap B'$.

(c) For every $B \in \mathcal{B}_0$ there exists $B' \in \mathcal{B}_0$ such that $B' + B' \subseteq B$.

Then $\mathcal{B}_0$ is a neighborhood basis at zero for a unique linear topology on X.

Proof. For every $x \in X$, set $\mathcal{B}_x = \mathcal{B}_0 + x$. We will first use Proposition 1.1.5 to show that the collections $\mathcal{B}_x$ for $x \in X$ give rise to a topology on X. Let us denote the properties from Proposition 1.1.5 by $(a') - (d')$.

(a'): Since $\mathcal{B}_0$ is assumed to be nonempty, it follows that each $\mathcal{B}_x$ is nonempty.

(b'): Every set in $\mathcal{B}_0$ is absorbing, hence contains 0. It follows that every set in $\mathcal{B}_x = \mathcal{B}_0 + x$ contains x.

(c'): If $B, B' \in \mathcal{B}_0$ then by (b) we can find $B'' \in \mathcal{B}_0$ such that $B'' \subseteq B \cap B'$. Hence $(B + x) \cap (B' + x) = B \cap B' + x$ for all $x \in X$.

(d'): Let $B \in \mathcal{B}_0$ and set

$$U = \{y \in B : \text{there exists } B' \in \mathcal{B}_0 \text{ such that } y + B' \subseteq B\}.$$

Then $0 \in U$ since $0 \in B$ and $0 + B = B$. Suppose $y \in U$, so that there exists some $B' \in \mathcal{B}_0$ such that $y + B' \subseteq B$. By (c) we can find $B'' \in \mathcal{B}_0$ such that $B'' + B'' \subseteq B'$. Hence $y + B'' \in \mathcal{B}_y$ has the property that $y + B'' \subseteq U$, since if $z \in y + B''$ then $z + B'' \subseteq y + B'' + B'' \subseteq y + B'$. This proves (d') for $x = 0$, and for general x we can do the same argument with $x + B$ and $x + U$ instead of B and U.

Proposition 1.1.5 now implies that the sets $\mathcal{B}_x$ are neighborhood bases for a (necessarily unique) topology on X. We must show that this topology is linear. First we show continuity of vector addition at $(x_0, y_0) \in X \times X$. Let $B \in \mathcal{B}_0$. By (c) from Proposition 2.2.5 we can find $B' \in \mathcal{B}_0$ such that $B' + B' \subseteq B$. If $x \in x_0 + B'$ and $y \in y_0 + B'$ then $x + y \in x_0 + y_0 + B' + B' \subseteq x_0 + y_0 + B$, so vector addition is continuous at (x_0, y_0).

Next we show continuity of scalar multiplication. Let $(\lambda_0, x_0) \in \mathbb{F} \times X$ and $B \in \mathcal{B}_0$. Pick $B' \in \mathcal{B}_0$ such that $B' + B' \subseteq B$. Since B' is absorbing we can find $\epsilon > 0$ such that $\lambda x_0 \in B'$ when $|\lambda| < \epsilon$. Now pick $n \in \mathbb{N}$ with $|\lambda_0| + \epsilon < n$. Thus, if $|\lambda - \lambda_0| < \epsilon$ then

$|\lambda/n| \leq (|\lambda_0|+\epsilon)/n < 1$. Assume that $|\lambda-\lambda_0| < \epsilon$ and $x \in x_0+(1/n)B'$. Since B' is balanced we have that

$$\lambda x = \lambda_0 x_0 + (\lambda-\lambda_0)x_0 + \lambda(x-x_0) \in \lambda_0 x_0 + B' + \frac{\lambda}{n}B' \subseteq \lambda_0 x_0 + B' + B' \subseteq \lambda_0 x_0 + B.$$

This shows that scalar multiplication is continuous. ∎

Definition 2.2.6. A linear topology on a vector space X is called *locally convex* if the zero vector has a neighborhood basis consisting of convex sets. A vector space equipped with a locally convex linear topology is called a *locally convex topological vector space.*

Remark. Note that every point in a locally convex topological vector space has a neighborhood basis consisting of convex sets, since if $\mathcal{B}_0$ is a neighborhood basis of 0 consisting of convex sets then $\mathcal{B}_x = x + \mathcal{B}_0$ is a neighborhood basis at x consisting of convex sets.

Proposition 2.2.7. *Let X be a locally convex topological vector space. Then X has a neighborhood basis at zero consisting of sets that are convex, balanced and absorbing. One can furthermore choose the neighborhoods to be open.*

Proof. Left as an exercise. ∎

Proposition 2.2.8. *Let X be a vector space and let $\mathcal{B}_0$ be a nonempty collection of subsets of X with the following properties:*

(a) Every $B \in \mathcal{B}_0$ is convex, balanced and absorbing.

(b) For every $B, B' \in \mathcal{B}_0$ there exists $B'' \in \mathcal{B}_0$ such that $B'' \subseteq B \cap B'$.

(c) For every $B \in \mathcal{B}_0$ there exists $\lambda \in (0, 1/2]$ such that $\lambda B \in \mathcal{B}_0$.

Then $\mathcal{B}_0$ is a neighborhood basis at zero for a unique linear topology on X, and this topology is locally convex.

Proof. To show that $\mathcal{B}_0$ is a neighborhood basis at zero for a unique linear topology on X, it suffices to show that (c) from Proposition 2.2.5 is satisfied. Let $B \in \mathcal{B}_0$. By (c) above we can find $\lambda \in (0, 1/2]$ such that $\lambda B \in \mathcal{B}_0$. Since B is convex we have that $\lambda B + \lambda B = 2\lambda B$. Since B is balanced and $2\lambda \leq 1$ we get $2\lambda B \subseteq B$.

It immediately follows that the topology determined by $\mathcal{B}_0$ is locally convex as the sets in $\mathcal{B}_0$ are convex by assumption. ∎

2.3 Topologies via semi-norms

In this section we look at semi-norms on topological vector spaces. The first result characterizes the continuity of a semi-norm in terms of its open and closed balls:

Proposition 2.3.1. *Let X be a topological vector space and let σ be a semi-norm on X. Then the following are equivalent:*

(a) σ is continuous.

(b) σ *is continuous at* 0.

(c) *The set* $B^\sigma = \{x \in X : \sigma(x) < 1\}$ *is open in* X.

(d) *The set* $\bar{B}^\sigma = \{x \in X : \sigma(x) \le 1\}$ *is a closed neighborhood of* 0.

Proof. Exercise. ∎

The following proposition concerns the continuity of gauges associated to sets in a topological vector space.

Proposition 2.3.2. *Let* X *be a topological vector space and let* C *be a convex and absorbing subset of* X. *Then the following hold:*

(a) *The gauge* m_C *of* C *is continuous if and only if* C *is a neighborhood of* 0.

(b) *We have that*

$$\operatorname{Int} C \subseteq B^{m_C} \subseteq C \subseteq \bar{B}^{m_C} \subseteq \operatorname{Cl} C.$$

Moreover, if C *is a neighborhood of* 0 *then* $\operatorname{Int} C = B^{m_C}$ *and* $\operatorname{Cl} C = \bar{B}^{m_C}$.

Proof. (a): Suppose first that C is a neighborhood of 0. Let $\epsilon > 0$. Then ϵC is a neighborhood of 0, and if $x \in \epsilon C$ then $m(x) = \inf\{\lambda > 0 : x \in \lambda C\} \le \epsilon$. Thus $m_C(\epsilon C) \subseteq \bar{B}_\epsilon(0)$ which shows that m_C is continuous at 0, hence everywhere by Proposition 2.3.1. Conversely, suppose m_C is continuous. Then $B^{m_C} = m_C^{-1}([0,1))$ is open and $B^{m_C} \subseteq C$ by Proposition 2.1.9, so C is a neighborhood of 0.

(b): Because of what we already know from Proposition 2.1.9 (d) it suffices to prove that $\operatorname{Int} C \subseteq B^{m_C}$ and $\bar{B}^{m_C} \subseteq \operatorname{Cl} C$. If $x \in \operatorname{Int} C$ then there exists an open set $U \subseteq C$ with $x \in U$. Thus we can find some $n \in \mathbb{N}$ such that $(1 + 1/n)x \in U \subseteq C$. But then $m_C((1+1/n)x) \le 1$ by Proposition 2.1.9 (d) so $m_C(x) \le 1/(1+1/n) = n/(n+1) < 1$. Next, suppose $x \in \bar{B}^{m_C}$. If $m_C(x) < 1$ then $x \in B^{m_C} \subseteq C \subseteq \operatorname{Cl} C$ so suppose $m_C(x) = 1$. Let U be an open neighborhood of x. By continuity of scalar multiplication we can find $\delta > 0$ such that $B_\delta(1)x \subseteq U$. Since $m_C(x) = 1$ we can find $\lambda > 1$ such that $x \in \lambda C$ and $|\lambda - 1| < \delta$. But then $|\lambda^{-1} - 1| = |1 - \lambda|/|\lambda| < \delta/|\lambda| < \delta$. Thus $\lambda^{-1} \in B_\delta(1)$ and since $\lambda^{-1}x \in C$ we have that $B_\delta(1)x \cap C \ne \emptyset$. Hence $U \cap C \ne \emptyset$. Since U was an arbitrary neighborhood of x we conclude that $x \in \operatorname{Cl} C$.

Finally, if C is a neighborhood of 0 then by (a) the gauge m_C is continuous. By Proposition 2.3.1 it follows that B^{m_C} is open and $\bar{B}^{m_C}$ is closed. By the inclusions $\operatorname{Int} C \subseteq B^{m_C} \subseteq C \subseteq \bar{B}^{m_C} \subseteq \operatorname{Cl} C$ it the follows that $\operatorname{Int} C = B^{m_C}$ and $\operatorname{Cl} C = \bar{B}^{m_C}$. ∎

Corollary 2.3.3. *In a topological vector space* X, *there is a one-to-one correspondence between (1) open, convex, balanced neighborhoods of* 0, *(2) closed, convex, balanced neighborhood of* 0 *and (3) continuous semi-norms on* X.

Proof. If C is an open (resp. closed), convex, absorbing, balanced subset of X then m_C is a semi-norm and $B^{m_C} = \operatorname{Int} C = C$ (resp. $\bar{B}^{m_C} = \operatorname{Cl} C = C$) by Proposition 2.3.2 (b). If σ is a continuous semi-norm on X then B^σ (resp. $\bar{B}^\sigma$) is an open (resp. closed), convex, balanced and absorbing set, and $m_{B^\sigma} = m_{\bar{B}^\sigma} = \sigma$ by Proposition 2.1.11. ∎

The following proposition will be our main way of obtaining locally convex topological vector spaces:

Proposition 2.3.4. *Let X be a vector space and let $(\sigma_i)_{i\in I}$ be a collection of semi-norms on X. For each finite subset I_0 of I and each $r > 0$ define*

$$N_{I_0,r} = \bigcap_{i\in I_0} B_r^{\sigma_i} = \{x \in X : \sigma_i(x) < r \text{ for each } i \in I_0\}. \tag{2.2}$$

Then the collection $\mathcal{B}_0 = \{N_{I_0,r} : I_0 \subseteq I \text{ finite}, r > 0\}$ defines an open neighborhood basis at zero for a locally convex linear topology on X. It is the initial topology determined by the family $(\sigma_i^y)_{(i,y)\in I\times X}$ where

$$\sigma_i^y(x) = \sigma_i(x - y) \quad \text{for } x \in X,\ i \in I \text{ and } y \in Y$$

and the weakest linear topology on X for which σ_i is continuous for every $i \in I$.

Proof. We must check that the axioms of Proposition 2.2.8 are satisfied. For (a), we know that each $B_r^{\sigma_i} = rB^{\sigma_i}$ is convex, balanced and absorbing from Proposition 2.1.7. Since $N_{I_0,r}$ is a finite intersection of such sets, it is also convex, balanced and absorbing. For (b), we note that $N_{I_0,r} \cap N_{I_1,s} \supseteq N_{I_0\cup I_1,\min\{r,s\}}$ for finite subsets I_0, I_1 of I and $r, s > 0$. For (c) we observe that $\lambda N_{I_0,r} = N_{I_0,\lambda r}$ for $\lambda > 0$. Thus we obtain a locally convex linear topology $\mathcal{T}$ on X such that $\mathcal{B}_0$ is a neighborhood basis at zero for $\mathcal{T}$.

We demonstrate that each set in $\mathcal{B}_0$ is open: If $x \in B^{\sigma_i}$ then let $\sigma_i(x) < r < 1$. Now $x + B_{1-r}^{\sigma_i}$ is a neighborhood of x and if $y \in B_{1-r}^{\sigma_i}$ then $\sigma_i(x + y) < r + (1 - r) = 1$. This shows that $x + B_{1-r}^{\sigma_i} \subseteq B^{\sigma_i}$, so B^{σ_i} is open. Since $N_{I_0,r}$ is a finite intersection of scalar multiples of B^{σ_i}'s, it is open as well.

Note that each σ_i is continuous in $\mathcal{T}$ by Proposition 2.3.1 since B^{σ_i} is open. Moreover each map σ_i^y is continuous, being the composition of the translation map $x \mapsto x - y$ with σ_i.

Suppose $\mathcal{T}'$ is any topology on X such that σ_i^y is continuous in $\mathcal{T}'$ for all $(i, y) \in I \times X$. Let $\mathcal{N}_y'$ be the neighborhood filter of $\mathcal{T}'$ at y. By continuity of the σ_i^y's, $\mathcal{N}_y'$ contains all sets of the form

$$y + N_{I_0,r} = \{x \in X : \sigma_i(x - y) < r \text{ for all } i \in I_0\} = \bigcap_{i\in I_0} (\sigma_i^y)^{-1}([0, r))$$

for $I_0 \subseteq I$ finite and $r > 0$. Thus $\mathcal{B}_y = y + \mathcal{B}_0 \subseteq \mathcal{N}_y'$, so since $\mathcal{B}_y$ is a neighborhood basis for $\mathcal{T}$ at y it follows that $\mathcal{T} \subseteq \mathcal{T}'$. Thus $\mathcal{T}$ is the initial topology on X determined by $(\sigma_i^y)_{(i,y)\in I\times X}$.

Finally, suppose $\mathcal{T}'$ is any linear topology on X such that σ_i is continuous in $\mathcal{T}'$ for every $i \in I$. Then every set in $\mathcal{B}_0$ is a neighborhood of zero in $\mathcal{T}'$, so we have $\mathcal{T} \subseteq \mathcal{T}'$ since linear topologies are determined by a neighborhood basis at zero. It follows that $\mathcal{T}$ is the weakest linear topology on X in which every σ_i is continuous. ■

Definition 2.3.5. Given a vector space X and a family $(\sigma_i)_{i\in I}$ of semi-norms on X, we call the topology constructed in Proposition 2.3.4 the *weak topology determined by* $(\sigma_i)_{i\in I}$.

Remark. Let X be a vector space equipped with the weak topology determined by a family of semi-norms $(\sigma_i)_{i\in I}$. Suppose $(x_\lambda)_{\lambda\in\Lambda}$ is a net in X and let $x \in X$. By Proposition 1.2.8 we have that $(x_\lambda)_{\lambda\in\Lambda} \to x$ if and only if $(\sigma_i^y(x_\lambda))_{\lambda\in\Lambda} \to \sigma_i^y(x)$ for all $y \in X$ and $i \in I$. For $y = x$ this gives $(\sigma_i(x_\lambda - x))_\lambda \to \sigma_i(0) = 0$ for all $i \in I$. On the other hand, if the convergence holds for $y = x$ then the reverse triangly inequality gives

$$|\sigma_i^y(x_\lambda) - \sigma_i^y(x)| = |\sigma_i(x_\lambda - y) - \sigma_i(x - y)| \leq \sigma_i(x_\lambda - y - (x - y)) = \sigma_i(x_\lambda - x) \to 0.$$

We conclude that $(x_\lambda)_\lambda \to x$ if and only if $(\sigma_i(x_\lambda - x))_\lambda \to 0$ for all $i \in I$.

Example 2.3.6. Let Ω be a topological space and let $X = C(\Omega)$ be the $\mathbb{F}$-vector space of continuous, $\mathbb{F}$-valued functions on Ω.

(a) Given $t \in \Omega$ we can define a semi-norm σ_t on X by $\sigma_t(f) = |f(t)|$. Then we can consider the weak topology induced by the family $(\sigma_t)_{t\in\Omega}$. A net $(f_\lambda)_{\lambda\in\Lambda}$ in X converges to f if and only if $(|f_\lambda(t) - f(t)|)_{\lambda\in\Lambda}$ converges to 0 for every $t \in \Omega$, i.e. $(f_\lambda(t))_{\lambda\in\Lambda} \to f(t)$ for all $t \in \Omega$. Thus we call this topology the *topology of pointwise convergence.*

(b) Given a compact set $K \subseteq \Omega$ we can define a semi-norm σ_K by $\sigma_K(f) = \sup_{t\in K} |f(t)|$. A net $(f_\lambda)_{\lambda\in\Lambda}$ converges to f in this topology if and only if $\sup_{t\in K} |f_\lambda(t) - f(t)|$ goes to zero for every compact set K in Ω. For this reason the topology determined by the family $(\sigma_K)_K$ is called the *topology of uniform convergence on compact subsets.*

Proposition 2.3.7. *Let X be a vector space with the weak topology induced by a family of semi-norms $(\sigma_i)_{i\in I}$. Then the following are equivalent:*

(a) X is Hausdorff.

(b) The family $(\sigma_i)_{i\in I}$ is separating, *i.e. whenever $x \in X$ and $\sigma_i(x) = 0$ for all $i \in I$ then $x = 0$.*

Proof. Note that $\sigma_i(x) = 0$ for all $i \in I$ if and only if $\sigma_i(x) < r$ for all $r > 0$ and $i \in I$ if and only if $x \in N_{I_0,r}$ for all finite subsets I_0 of I and $r > 0$. Since the latter family is a neighborhood basis at zero we have that $\sigma_i(x) = 0$ for all $i \in I$ if and only if x is in every neighborhood of zero. By an exercise we have that X is Hausdorff if and only if $x \in N$ for all $N \in \mathcal{N}_0$ implies $x = 0$. Hence X is Hausdorff if and only if $(\sigma_i)_{i\in I}$ is separating. ■

The following proposition tells us that every locally convex linear topology on a vector space comes from a family of semi-norms. Thus, defining a topology on a vector space via a neighborhood basis of convex sets or via semi-norms are equivalent.

Proposition 2.3.8. *Every locally convex linear topology $\mathcal{T}$ on a vector space X is induced by a family of semi-norms. In fact, if $\mathcal{B}_0$ is a neighborhood basis at zero for $\mathcal{T}$ consisting of open, convex, balanced and absorbing sets as in Proposition 2.2.7, then $\mathcal{T}$ is the weak topology induced by the family $(m_U)_{U\in\mathcal{B}_0}$ of semi-norms.*

Proof. The topology on X induced by $(m_U)_{U\in\mathcal{B}_0}$ has neighborhood basis at zero $\mathcal{B}_0'$ consisting of the sets $N_{I_0,r} = \{x \in X : m_U(x) < r \text{ for all } B \in I_0\}$ where r ranges over positive numbers and I_0 ranges over finite subsets of $\mathcal{B}_0$. We must show that $\mathcal{B}_0$ and $\mathcal{B}_0'$ determine the same topology. By Proposition 2.3.2 we have that

$$N_{I_0,r} = \bigcap_{U\in I_0} B_r^{m_U} = \bigcap_{U\in I_0} rB^{m_U} = r \bigcap_{U\in I_0} U.$$

for finite $I_0 \subseteq \mathcal{B}_0$ and $r > 0$. By setting $I_0 = \{U\}$ and $r = 1$ we have that $N_{I_0,r} = U$, which shows that $\mathcal{B}_0 \subseteq \mathcal{B}_0'$. Conversely, since every set of the form $r\bigcap_{U\in I_0} U$ is a neighborhood of 0 in the topology on X determined by $\mathcal{B}_0$, there exists a neighborhood $U' \in \mathcal{B}_0$ such that $U' \subseteq r\bigcap_{U\in I_0} U$. Hence $\mathcal{B}_0$ is a refinement of $\mathcal{B}_0'$, so the two topologies are equal. ■

2.4 The Hahn–Banach separation theorems

If $T\colon X \to Y$ is a linear map between normed spaces, continuity of T is equivalent to T being *bounded*, that is, there exists $K \geq 0$ such that $\|T(x)\| \leq K\|x\|$ for all $x \in X$. The following proposition is a generalization of this characterization of continuity to the setting of a linear map between locally convex topological vector spaces.

Proposition 2.4.1. *Let X and Y be locally convex topological vector spaces. Suppose the topology of X (resp. Y) is induced by a family of semi-norms $(\sigma_i)_{i\in I}$ on X (resp. $(\rho_j)_{j\in J}$ on Y). Then the following are equivalent:*

(a) T is continuous.

(b) T is continuous at 0.

(c) For every continuous semi-norm ρ on Y there exists a continuous semi-norm σ on X such that
$$\rho(T(x)) \leq \sigma(x) \quad \text{for all } x \in X.$$

(d) For every $j \in J$ there exists a finite subset I_0 of I and $K \geq 0$ such that
$$\rho_j(T(x)) \leq K \max_{i\in I_0} \sigma_i(x) \quad \text{for all } x \in X.$$

Proof. $(a) \Leftrightarrow (b)$: Exactly the same proof as for normed spaces.

$(a) \Rightarrow (c)$: If T is continuous and ρ is a continuous semi-norm on Y then $\sigma(x) = \rho(T(x))$ is a continuous semi-norm on X.

$(c) \Rightarrow (d)$: Let $j \in J$. Then by (c) there exists a continuous semi-norm σ on X such that $\rho_j(T(x)) \leq \sigma(x)$ for all $x \in X$. Since σ is continuous it follows from Proposition 2.3.1 that B^σ is an open neighborhood of 0. Thus, by Proposition 2.3.4, there exists a finite subset I_0 of I and $r > 0$ such that $N_{I_0,r} \subseteq B^\sigma$. Let $x \in X$ and set $m = \max_{i\in I_0} \sigma_i(x)$. Then for all $i \in I_0$ we have
$$\sigma_i((r/2)m^{-1}x) = (r/2)\sigma_i(x)/m \leq r/2 < r$$
so $\sigma(((r/2)m^{-1}x) < 1$, or, $\sigma(x) \leq K \max_{i\in I_0} \sigma_i(x)$ where $K = r/2$. Combined with $\rho_j(T(x)) \leq \sigma(x)$ we have arrived at (d).

$(d) \Rightarrow (b)$. Assume that (d) holds. Let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in X that converges to 0. Then $(\sigma_i(x_\lambda))_{\lambda\in\Lambda}$ converges to 0 for every $i \in I$. Let $j \in J$. By the assumption, the net $(\rho_j(T(x_\lambda)))_\lambda$ converges to 0. Since j was arbitrary, we conclude by Proposition 2.3.4 that $(T(x_\lambda))_{\lambda\in\Lambda}$ converges to 0. Exercise. ∎

Corollary 2.4.2. *Let $\phi\colon X \to \mathbb{F}$ be a linear functional on a topological vector space. Then ϕ is continuous if and only if there exists a continuous semi-norm σ on X such that $|\phi(x)| \leq \sigma(x)$ for all $x \in X$.*

Proof. If ϕ is continuous then $\sigma = |\phi|$ does the trick. Conversely, suppose there exists a semi-norm σ on X such that $|\phi(x)| \leq \sigma(x)$ for all $x \in X$. Every semi-norm on $\mathbb{F}$ is a nonnegative multiple of the absolute value on $\mathbb{F}$ (prove this!). For every $r \geq 0$ we have that $r|\phi(x)| \leq r\sigma(x)$ so (c) of Proposition 2.4.1 is verified. Thus ϕ is continuous. ∎

Recall the following fundamental result from MAT4410:

Theorem 2.4.3 (Hahn–Banach Theorem)**.** *Let X be a real vector space and let p be a sublinear functional on X. Let X_0 be a linear subspace of X and let ϕ_0 be a linear functional on X_0. Suppose that*

$$\phi_0(x) \leq p(x) \quad \text{for all } x \in X_0.$$

Then there exists a linear functional ϕ on X such that $\phi|_{X_0} = \phi_0$ and

$$\phi(x) \leq p(x) \quad \text{for all } x \in X.$$

Let X be a vector space over $\mathbb{R}$. By a *hyperplane* in X we mean a set of the form $H = \phi^{-1}(c)$ for some linear functional on X and some $c \in \mathbb{R}$. A hyperplane defines two half-spaces

$$L = \{x \in X : \phi(x) \leq c\}, \qquad U = \{x \in X : \phi(x) \geq c\}.$$

Given two disjoint subsets A and B of X, we are interested in whether we can separate the two by a hyperplane: That is, to which extent we can find a hyperplane H such that A sits inside L and B sits inside U as above. We will consider the following two degress of separation:

(a) Strict separation: We can find a linear functional ϕ on X such that $\phi(x) < c < \phi(y)$ for all $x \in A$ and $y \in B$.

(b) Strong separation: We can find a linear functional ϕ on X and $r > 0$ such that $\phi(x) < c - r < c + r < \phi(y)$ for all $x \in A$ and $y \in B$. Equivalently $\sup_{x \in A} \phi(x) < \inf_{y \in B} \phi(y)$.

For complex vector spaces we can ask the same questions if we consider only the real part of a linear functional.

The following important theorem gives conditions on A and B in a topological vector space that guarantee both strict separation and strong separation:

Theorem 2.4.4 (Hahn–Banach Separation Theorem)**.** *Let X be a topological vector space and let A and B be nonempty, convex subsets of X with $A \cap B = \emptyset$.*

(a) Suppose A is open. Then there exists a continuous linear functional ϕ on X and $c \in \mathbb{R}$ such that

$$\operatorname{Re} \phi(x) < c \leq \operatorname{Re} \phi(y) \quad \text{for all } x \in A \text{ and } y \in B.$$

If B is open as well, then both of the inequalities above are strict.

(b) Suppose X is locally convex and suppose that A is closed and B is compact. Then there exists a continuous linear functional ϕ on X and $c, r \in \mathbb{R}$ such that

$$\operatorname{Re} \phi(x) \leq c - r < c + r \leq \operatorname{Re} \phi(y) \quad \text{for all } x \in A \text{ and } y \in B.$$

To prove Theorem 2.4.4 we need a couple of following lemmas as preparation.

Lemma 2.4.5. *Every nontrivial linear functional ϕ on a topological vector space X is an open map.*

Proof. Let $U \subseteq X$ be open and nonempty. Since ϕ is nontrivial we can find $x_0 \in X$ with $\phi(x_0) = 1$. Let $\mu \in \phi(U)$ and pick some $x \in U$ with $\phi(x) = \mu$. Then $U - x$ is absorbing by Proposition 2.2.3 (b) so we can find $\epsilon > 0$ such that $\lambda x_0 \in U - x$ when $|\lambda| < \epsilon$. Thus $x + \lambda x_0 \in U$ so $\mu + \lambda = \phi(x + \lambda x_0) \in \phi(U)$ when $|\lambda| < \epsilon$. This shows that $B_\epsilon(\mu) \subseteq \phi(U)$, so $\phi(U)$ is open. ■

Lemma 2.4.6. *Let X be a topological vector space. Let U be an open, convex subset of X that does not contain* 0*. Then there exists a continuous linear functional ϕ on X such that* $\operatorname{Re}\phi$ *is positive on U.*

Proof. Assume first that $\mathbb{F} = \mathbb{R}$. Let $x_0 \in -U$. Then $U + x_0$ is a convex, open neighborhood of 0 so by Proposition 2.3.2 (b) we can write we can write $U + x_0 = B^p$ where p is the gauge of $U + x_0$. Since $0 \notin U$ we have that $x_0 \neq 0$ and $x_0 \notin x_0 + U$, so $p(x_0) \geq 1$. Set $X_0 = \operatorname{span}\{x_0\}$ and define $\phi_0\colon X_0 \to \mathbb{R}$ by $\phi_0(\lambda x_0) = \lambda$. Then ϕ_0 is a linear functional on X_0. If $\lambda \geq 0$ then $\phi_0(\lambda x_0) = \lambda \leq \lambda p(x_0) = p(\lambda x_0)$. If $\lambda < 0$ then $\phi_0(\lambda x_0) = \lambda < 0 \leq p(\lambda x_0)$. In any case p is a sublinear functional that dominates ϕ_0 on X_0, so by Theorem 2.4.3 there exists an extension ϕ of ϕ_0 to X with $\phi(x) \leq p(x)$ for all $x \in X$. Now $\sigma(x) = \max\{p(x), p(-x)\}$ is a semi-norm and $|\phi(x)| \leq \sigma(x)$ for all $x \in X$, so by Corollary 2.4.2 ϕ is continuous.

Now if $\phi(x) = 0$ then $1 = \phi(x) + \phi(x_0) = \phi(x + x_0) \leq p(x + x_0)$. Thus $x + x_0 \notin U + x_0$, so $x \notin U$. This shows that if $x \in U$ then either $\phi(x) > 0$ or $\phi(x) < 0$. Since U is connected it follows that ϕ is either positive or negative on the whole of U. Thus, possibly interchanging ϕ with $-\phi$, we have a linear functional as desired.

Let now $\mathbb{F} = \mathbb{C}$. We can still view X as a real vector space, and the the notion of convexity does not change. Thus, by what we already proved we can find a continuous $\mathbb{R}$-linear functional $\psi\colon X \to \mathbb{R}$ such that ψ is positive on U. Define $\phi\colon X \to \mathbb{C}$ by

$$\phi(x) = \psi(x) - i\psi(ix) \quad \text{for } x \in \mathbb{C}.$$

Then ϕ is obviously continuous and $\mathbb{R}$-linear, but it is also $\mathbb{C}$-linear, as

$$\phi(ix) = \psi(ix) - i\psi(i^2 x) = \psi(ix) + i\psi(x) = i(\psi(x) - i\psi(ix)) = i\phi(x).$$

Since $\operatorname{Re}\phi = \psi$ we have $\operatorname{Re}\phi > 0$ on U. ■

Proof of Theorem 2.4.4. (a): Set $U = B - A$. Then U is open ($U = \bigcup_{y \in B}(b - A)$) and convex (Proposition 2.1.4). Moreover, $0 \notin U$ since $A \cap B = \emptyset$. By Lemma 2.4.6 there exists a continuous linear functional ϕ on X such that $\operatorname{Re}\phi$ is positive on $B - A$. Hence $\operatorname{Re}\phi(x) < \operatorname{Re}\phi(y)$ for all $x \in A$ and $y \in B$. Thus $\operatorname{Re}\phi(A)$ and $\operatorname{Re}\phi(B)$ are connected, disjoint subsets of $\mathbb{R}$. If A is open then ϕ is open by Lemma 2.4.5, so $\operatorname{Re}\phi$ must be open ($\operatorname{Re}\colon \mathbb{C} \to \mathbb{R}$ is a coordinate projection, hence an open map). It follows that $\operatorname{Re}\phi(A)$ is an open interval, thus bounded away from its right endpoint c. This gives $\operatorname{Re}\phi(x) < c \leq \operatorname{Re}\phi(y)$ for all $x \in A$ and $y \in B$. If B is open as well then we get strict inequalities for the same reason.

(b): By the assumptions, B is a compact subset of the open set A^c, so we can find a convex open neighborhood V of zero such that $B + V \subseteq A^c$ (exercise!). Now $B + V$ and A are nonempty, convex sets that are disjoint and $B + V$ is open. By (a) we can find a continuous linear functional ϕ on X and $c \in \mathbb{R}$ such that $\operatorname{Re}\phi(x) \leq c < \operatorname{Re}\phi(y)$ for all $x \in A$ and $y \in B + V$. Again, since V is open neighborhood of 0 and $\operatorname{Re}\phi$ is an open map, we can fit an interval of the form $(-r, r)$ for some $r > 0$ inside $\operatorname{Re}\phi(V)$. Consequently there exists $x_0 \in \operatorname{Re}\phi(V)$ with $\operatorname{Re}\phi(x_0) < 0$. It follows that for every $x \in A$ and $y \in B$ we have

$$\operatorname{Re}\phi(x) \leq c < \operatorname{Re}\phi(y + x_0) = \operatorname{Re}\phi(y) + \operatorname{Re}\phi(x_0).$$

Thus we have that $\sup_{x\in A} \operatorname{Re}\phi(x) \leq \inf_{y\in B} \operatorname{Re}\phi(y) + \operatorname{Re}\phi(x_0) < \inf_{y\in B} \operatorname{Re}\phi(y)$. ■

2.5 The weak and weak* topologies

Given a topological vector space X, we denote by X^* the (continuous) dual space, that is, the vector space of continuous linear functionals on X.

Given $\phi \in X^*$, we have already noted that the map $|\phi|\colon X \to [0, \infty)$ given by $|\phi|(x) = |\phi(x)|$ for $x \in X$ is a semi-norm on X.

Definition 2.5.1. Let X be a topological vector space. The topology on X determined by the family $(|\phi|)_{\phi\in X^*}$ of semi-norms on X is called the *weak* topology on X.

Note that by Proposition 2.3.4 we have that a net $(x_\lambda)_{\lambda\in\Lambda}$ in X converges to x in the weak topology if and only if $(\phi(x_\lambda))_{\lambda\in\Lambda}$ converges to $\phi(x)$ for every $\phi \in X^*$. Therefore the weak topology on X is weaker than the original topology on X.

Observation 2.5.2. Suppose X is a nontrivial Hausdorff locally convex topological vector space. Let $x \in X$ be nonzero. Then $\{0\}$ and $\{x\}$ are nonempy, convex, compact subsets of X that are disjoint, so by Theorem 2.4.4 we can find $\phi \in X^*$ such that $0 = \operatorname{Re}\phi(0) < \operatorname{Re}\phi(x)$. In particular $\phi(x) \neq 0$. This shows not only that the dual space X^* is nontrivial, but also that the family $(|\phi|)_{\phi\in X^*}$ of semi-norms is separating. Consequently the weak topology on X is Hausdorff by Proposition 2.3.7.

The following example shows that the weak topology might be strictly weaker than the original topology on a topological vector space.

Example 2.5.3. Let H be a Hilbert space. Then by the Riesz' representation theorem, every linear functional on H is of the form $x \mapsto \langle x, y\rangle$ for some $y \in H$. Hence, a sequence $(x_n)_{n\in\mathbb{N}}$ in H converges weakly to x if and only if $(\langle x_n, y\rangle)_n$ converges to $\langle x, y\rangle$ for every $y \in H$. Let us look at a particular example. Suppose H is infinite-dimensional and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in H. Then for $m, n \in \mathbb{N}$ with $m \neq n$ we have

$$\|e_m - e_n\|^2 = \|e_m\|^2 - 2\operatorname{Re}\langle e_m, e_n\rangle + \|e_n\|^2 = 2.$$

This shows that $(e_n)_{n\in\mathbb{N}}$ does not converge in norm. However, if $y \in H$ then by Bessel's inequality we have that $\sum_n |\langle y, e_n\rangle|^2 \leq \|y\|^2$ so the sequence $(|\langle y, e_n\rangle|^2)_n$ must go to zero for every $y \in H$. Thus $(e_n)_n$ converges weakly to 0.

In fact, one can show that the weak topology on X never agrees with the original topology on X when X is an infinite-dimensional locally convex topological vector space.

Proposition 2.5.4. *Let C be a convex subset of a locally convex topological vector space X. Then C is closed if and only if it is weakly closed.*

Proof. Since the weak topology on X is weaker than the original topology on X, weakly closed sets are closed (in the original topology). Suppose that C is closed. Let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in C that converges weakly to $x \in X$. Assume for a contradiction that $x \notin C$. Since C^c is an open neighborhood of x and X is locally convex we can find an open convex neighborhood $U \subseteq C^c$ of x. By Theorem 2.4.4 (a) we can find $\phi \in X^*$ such that $\operatorname{Re}\phi(x') < c \leq \operatorname{Re}\phi(y)$ for all $x' \in U$ and $y \in C$. In particular $\operatorname{Re}\phi(x) < \inf\{\phi(y) : y \in C\}$. But then $\operatorname{Re}\phi(x) < \inf\{\operatorname{Re}\phi(x_\lambda) : \lambda \in \Lambda\}$ which contradicts that $(\phi(x_\lambda))_\lambda \to \phi(x)$. Hence C is weakly closed. ■

When X is a normed space we can always equip X^* with the operator norm, turning it into a normed space as well. When X is only a topological vector space, the operator norm is lacking. However, we can always equip X^* with the following topology:

Definition 2.5.5. Let X be a topological vector space. For each $x \in X$ let σ_x denote the semi-norm on the dual space X^* given by

$$\sigma_x(\phi) = |\phi(x)| \quad \text{for } \phi \in X^*.$$

The topology on X^* determined by the family $(\sigma_x)_{x\in X}$ is called the *weak* topology* on X^*.

By Proposition 2.3.4 a net $(\phi_\lambda)_{\lambda\in\Lambda}$ in X^* converges weak*ly to ϕ if and only if for every $x \in X$ the net $(\phi_\lambda(x))_{\lambda\in\Lambda}$ converges to $\phi(x)$. Thus, the weak* topology on X^* is the topology of pointwise convergence of continuous linear functionals.

Example 2.5.6. Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space and consider the Banach space $X = L^1(\Omega, \mathcal{A}, \mu)$. Then X^* is isomorphic to $L^\infty(\Omega, \mathcal{A}, \mu)$, where a function $g \in L^\infty$ is mapped to the linear functional on X given by $f \mapsto \int_\Omega fg \,\mathrm{d}\mu$. A net $(g_\lambda)_{\lambda\in\Lambda}$ of functions in L^∞ converges to g in the weak* topology if and only if $\int_\Omega fg_\lambda \,\mathrm{d}\mu \to \int_\Omega fg \,\mathrm{d}\mu$ for all $f \in L^1$.

Theorem 2.5.7 (The Banach–Alaoglu Theorem)**.** *Let X be a normed and space and denote by $\bar{B}^*$ the closed unit ball of X^*, i.e.*

$$\bar{B}^* = \{\phi \in X^* : \|\phi\| \leq 1\}.$$

Then $\bar{B}^*$ is compact in the weak topology.*

Proof. Let K be the product space $K = \prod_{x\in X} B_{\|x\|}(0)$ where $B_{\|x\|}(0)$ is the closed ball of radius $\|x\|$ centered at 0 in $\mathbb{F}$. Each of the sets $B_{\|x\|}(0)$ is a closed and bounded subset of $\mathbb{F}$, hence compact by the Heine–Borel Theorem. It follows that the product space K is compact by Tychonoff's Theorem (Theorem 1.3.8). An element of K is a function $\phi\colon X \to \mathbb{F}$ such that $|\phi(x)| \leq \|x\|$ for every $x \in X$. Thus, we can identify $\bar{B}^*$ with the subset of K consisting of linear functions. The subspace topology on $\bar{B}^*$ coming from the product topology on K is exactly the weak* topology. Hence, to show that $\bar{B}^*$ is compact, it suffices to show that $\bar{B}^*$ is closed in K. Let $(\phi_\gamma)_{\gamma\in\Gamma}$ be a net in $\bar{B}^*$ that converges to $\phi \in K$. Letting $x, y \in X$ and $\lambda, \mu \in \mathbb{F}$, we have that

$$\phi(\lambda x + \mu y) = \lim_\gamma \phi_\gamma(\lambda x + \mu y) = \lambda \lim_\gamma \phi_\gamma(x) + \mu \lim_\gamma \phi_\gamma(y) = \lambda\phi(x) + \mu\phi(y).$$

This shows that $\bar{B}^*$ is closed in K which finishes the proof. ■

Let X be a normed space. Denote by X^* its dual space, which is again a normed space with respect to the operator norm $\|\phi\| = \sup_{\|x\|=1} |\phi(x)|$ for $\phi \in X^*$. Considering X^* as a nomed vector space in this way, we can form its continuous dual $X^{**} = (X^*)^*$, which is again a normed space. This space is called the *bidual* of X.

Proposition 2.5.8. *Let X be a normed space. Then for every $x \in X$ the evaluation functional $\mathrm{ev}_x \colon X^* \to \mathbb{F}$ given by $\mathrm{ev}_x(\phi) = \phi(x)$ is an element of X^{**}. Furthermore, the map $\Phi \colon X \to X^{**}$ by*

$$\Phi(x) = \mathrm{ev}_x \quad \textit{for } x \in X$$

is a linear isometry.

Proof. Let $x \in X$, $\phi, \psi \in X^*$ and $\lambda, \mu \in \mathbb{F}$. Then

$$\mathrm{ev}_x(\lambda\phi + \mu\psi) = (\lambda\phi + \mu\psi)(x) = \lambda\phi(x) + \mu\psi(x) = \lambda\,\mathrm{ev}_x(\phi) + \mu\,\mathrm{ev}_x(\psi)$$

shows that ev_x is a linear functional on X^*. Furthermore

$$|\,\mathrm{ev}_x(\phi)| = |\phi(x)| \leq \|\phi\|\|x\|$$

which shows that ev_x is bounded, with $\|\,\mathrm{ev}_x\,\| \leq \|x\|$. By a well-known corollary of the Hahn–Banach theorem there exists $\phi_0 \in X^*$ with $\phi_0(x) = \|x\|$ so we actually have that

$$\|\,\mathrm{ev}_x\,\| = \sup_{\|\phi\|=1} |\phi(x)| \geq |\phi_0(x)| = 1$$

so $\|\,\mathrm{ev}_x\,\| = \|x\|$. Now let $x, y \in X$, $\lambda, \mu \in \mathbb{F}$ and $\phi \in X^*$. Using the linearity of ϕ we have that

$$\Phi(\lambda x + \mu y)(\phi) = \phi(\lambda x + \mu y) = \lambda\phi(x) + \mu\phi(y) = \lambda\Phi(x)(\phi) + \mu\Phi(y)(\phi).$$

Finally $\|\Phi(x)\| = \|\,\mathrm{ev}_x\,\| = \|x\|$ which shows that Φ is an isometry. ∎

Note that Φ is injective since it is an isometry, but not surjective in general.

Definition 2.5.9. A normed space X is called *reflexive* if the isometry $\Phi \colon X \to X^{**}$ in Proposition 2.5.8 is surjective.

Example 2.5.10. (a) For any normed space X we know that the dual space X^* is a Banach space even if X is not. Thus, the bidual X^{**} of any normed space is a Banach space, so for X to be reflexive, X must be a Banach space.

(b) Hilbert spaces are reflexive: Recall that for a Hilbert space H, every $\phi \in H^*$ is of the form ϕ_y for some unique $y \in H$ where $\phi_y(x) = \langle x, y\rangle$ for $x \in H$. Moreover, $\langle \phi_y, \phi_z\rangle = \langle z, y\rangle$ defines an inner product on H^*. If $\alpha \in H^{**}$ we can for the same reason find $\psi \in H^*$ such that $\alpha(\phi) = \langle \phi, \psi\rangle$ for $\phi \in H^*$, and we must have $\psi = \phi_z$ for some $z \in H$. But then

$$\Phi(z)(\phi_y) = \phi_y(z) = \langle z, y\rangle = \langle \phi_y, \phi_z\rangle = \alpha(\phi_y)$$

for all $y \in H$. Hence $\Phi(z) = \alpha$ which shows that Φ is surjective.

(c) Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space, let $1 < p < \infty$ and set $X = L^p(X, \mathcal{A}, \mu)$. Then X is reflexive since $X^* \cong L^q$ where $1/p + 1/q = 1$, and repeating the procedure gives $X^{**} \cong L^p$. As in the previous example one can check that this isomorphism is given by Φ. However, if $p = 1$ then the bidual of L^1 can be in general much bigger than L^1. Hence L^1 is not reflexive in general.

2.6 The Krein–Milman theorem

Let x and y be points in a vector space X. The *open line segment* between x and y is the set

$$(x, y) = \{(1-t)x + ty : 0 < t < 1\}.$$

We say that $z \in X$ lies *strictly between* x and y if $z \in (x, y)$.

Definition 2.6.1. Let C be a nonempty convex subset of a vector space X. We say that a nonempty convex subset F of C is a *face* of C if whenever $x, y \in C$ and some point in F lies strictly between x and y, then $x, y \in F$. That is, whenever $x, y \in C$ and $(1-t)x + ty \in F$ for some $t \in (0, 1)$, then $x, y \in F$.

Notice that if F', F and C are convex sets such that F' is a face of F and F is a face of C then F' is also a face of C.

Definition 2.6.2. Let X be a vector space and let C be a convex subset of X. We say that $y \in C$ is an *extreme point* for C if y does not lie strictly between any two distinct points in C. That is, whenever $y = (1-t)x_1 + tx_2$ for some $t \in (0, 1)$ and $x_1, x_2 \in C$ then $x_1 = x_2$. We denote by $\mathrm{ex}(C)$ the set of extreme points of C.

Note that an extreme point for C is precisely the same as a singleton face: That is, $\{x\}$ is a face for C if and only if x is an extreme point for C.

Example 2.6.3. Let X be a nonzero normed space and suppose x is an extreme point for the closed unit ball $\bar{B}$ of X. Then $\|x\| = 1$, which we can see using a proof by contradiction: If $x = 0$ then x lies strictly between $-y$ and y for any $y \in X$ with $\|y\| = 1$. If $0 < \|x\| < 1$ then x lies strictly between 0 and $x/\|x\|$. Thus the only option is that $\|x\| = 1$.

The converse does not hold in general (take e.g. $\mathbb{R}^2$ with the ∞-norm) , although it holds e.g. when X is a Hilbert space. To see this, suppose $\|x\| = 1$ and assume that $x = (1-t)x_1 + tx_2$ for $0 < t < 1$, and $\|x_1\| \leq 1$, $\|x_2\| \leq 1$. Then

$$1 = \|x\| \leq (1-t)\|x_1\| + t\|x_2\|.$$

If $\|x_1\| < 1$ then we get $1 < (1-t) \cdot 1 + t\|x_2\| \leq (1-t) + t \cdot 1 = 1$, a contradiction, so $\|x_1\| = 1$. Similarly we also need $\|x_2\| = 1$. Now

$$1 = \|x\|^2 = (1-t)^2\|x_1\|^2 + t^2\|x_2\|^2 + 2(1-t)t\,\mathrm{Re}\langle x_1, x_2\rangle = (1-t)^2 + t^2 + 2(1-t)t\,\mathrm{Re}\langle x_1, x_2\rangle.$$

Solving for $\mathrm{Re}\langle x_1, x_2\rangle$ we obtain

$$\mathrm{Re}\langle x_1, x_2\rangle = \frac{1 - (1-t)^2 - t^2}{t(1-t)} = 1.$$

Thus $1 = \mathrm{Re}\langle x_1, x_2\rangle \leq |\langle x_1, x_2\rangle| \leq \|x_1\|\|x_2\| = 1$. Thus we have equality in the Cauchy–Schwarz inequality for x_1 and x_2 which means that they are parallel, say $x_1 = cx_2$. But then $1 = \|x_1\| = |c|\|x_2\| = |c|$, and by the above, $1 = \mathrm{Re}\langle x_1, x_2\rangle = \mathrm{Re}(c\|x_2\|^2) = \mathrm{Re}\, c$, so we must have $c = 1$. Thus $x_1 = x_2$. This shows that x is an extreme point for $\bar{B}$.

The present section is dedicated to proving and exploring the consequences of the following theorem:

Theorem 2.6.4 (The Krein–Milman Theorem)**.** *Let X be a Hausdorff locally convex topological vector space. Let K be a compact convex subset of X. Then*

$$\mathrm{Cl}(\mathrm{co}(\mathrm{ex}(K))) = K.$$

Before we present the proof we need some preparation.

Lemma 2.6.5. *Let K be a nonempty compact convex subset of a topological vector space X. Let $\phi \in X^*$ and set $m = \inf\{\operatorname{Re}\phi(x) : x \in K\}$. Then $F = \{x \in K : \operatorname{Re}\phi(x) = m\}$ is a compact face of K.*

Proof. Let $f = \operatorname{Re}\phi\colon K \to \mathbb{R}$. Then the continuity of f implies that $f(K)$ is a compact subset of $\mathbb{R}$, so m is a well-defined real number and f attains m at least once. Thus F is nonempty. Moreover F is convex (being the inverse image under a linear map of a point) and closed ($F = f^{-1}(\{m\})$), hence compact.

Finally we show that F is a face of K: Suppose $x, y \in K$ and $(1-t)x + ty \in F$ for some $0 < t < 1$. Then $m = f((1-t)x + ty) = (1-t)f(x) + tf(y)$. Since $f(x) \geq m$ and $f(y) \geq m$ this implies $f(x) = f(y) = m$, so $x, y \in F$. ■

An important part of the proof of the Krein–Milman Theorem involves showing that a nonempty compact convex set has at least one extreme point. We present this argument in the following separate lemma:

Lemma 2.6.6. *Let K be a nonempty compact convex subset of a Hausdorff locally convex topological vector space. Then* $\mathrm{ex}(K) \neq \emptyset$.

Proof. Let $\mathcal{F}$ be the set of compact faces of K. Then $\mathcal{F} \neq \emptyset$ since $K \in \mathcal{F}$. We equip $\mathcal{F}$ with the partial order of reverse inclusion, and our goal is to employ Zorn's lemma. Assume that $\mathcal{C}$ is a nonempty chain in $\mathcal{F}$. Whenever $\mathcal{C}_0$ is a finite subset of $\mathcal{C}$ then $\bigcap_{F \in \mathcal{C}_0} F$ is nonempty. By compactness of K we have that $F_0 := \bigcap_{F \in \mathcal{C}} F \neq \emptyset$. Thus F_0 is a face of K and F_0 is compact, being an intersection of compact sets. This means that $F_0 \in \mathcal{F}$ and since $F_0 \subseteq F$ for all $F \in \mathcal{C}$ Zorn's lemma gives us the existence of a maximal element E of $\mathcal{F}$.

We claim that E must be a singleton. Suppose otherwise, i.e. that we can find $x_0, y_0 \in E$ with $x_0 \neq y_0$. Since X is Hausdorff $\{x_0\}$ and $\{y_0\}$ are closed (and compact), so by Theorem 2.4.4 (*b*) we can find $\phi \in X^*$ such that $\operatorname{Re}\phi(x_0) < \operatorname{Re}\phi(y_0)$. By Lemma 2.6.5 we can find $m \in \mathbb{R}$ such that $F = \{x \in E : \operatorname{Re}\phi(x) = m\}$ is a compact face of E. Thus F is a compact face of K so $F \in \mathcal{F}$. Since $F \subseteq E$, maximality of E gives $F = E$. But then $x_0, y_0 \in F$ so $\operatorname{Re}\phi(x_0) = m = \operatorname{Re}\phi(y_0)$, a contradiction. Hence E is singleton, and its element must be an extreme point of K. ■

We are now ready to prove the Krein–Milman Theorem:

Proof of Theorem 2.6.4. If $K = \emptyset$ there is nothing to prove, so assume that $K \neq \emptyset$. Note that since $\mathrm{ex}(K) \subseteq K$ we have that $\mathrm{co}(\mathrm{ex}(K)) \subseteq K$ (since K is convex) and thus $\mathrm{Cl}(\mathrm{co}(\mathrm{ex}(K))) \subseteq K$ (since K is closed). It remains to show the reverse inclusion.

Suppose for a contradiction that there exists $x_0 \in K$ with $x_0 \notin \mathrm{Cl}(\mathrm{co}(\mathrm{ex}(K)))$. By Theorem 2.4.4 (*b*) we can find $\phi \in X^*$ and $t \in \mathbb{R}$ such that

$$\operatorname{Re}\phi(x_0) < t < \operatorname{Re}\phi(y) \quad \text{for all } y \in \mathrm{Cl}(\mathrm{co}(\mathrm{ex}(K))).$$

Set $m = \inf\{\operatorname{Re}\phi(x) : x \in K\}$ and $F = \{x \in K : \operatorname{Re}\phi(x) = m\}$. By Lemma 2.6.5 F is a compact face of K. Since $\operatorname{Re}\phi(x_0) \geq m$, the above inequality implies that $F \cap \operatorname{Cl}(\operatorname{co}(\operatorname{ex}(K))) = \emptyset$. Since F is a nonempty compact convex subset of K, it has an extreme point x_1 by Lemma 2.6.6. Since F is a face of K it follows that x_1 is an extreme point for K as well. But then $x_1 \in F \cap \operatorname{Cl}(\operatorname{co}(\operatorname{ex}(K))$ which is a contradiction. We conclude that $K = \operatorname{Cl}(\operatorname{co}(\operatorname{ex}(K)))$ which finishes the proof.

■

Chapter 3

Banach algebras

3.1 Basic definitions and examples

For the moment, $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$, although we will eventually work exlusively over the complex numbers.

Definition 3.1.1. An $\mathbb{F}$-*algebra* is an $\mathbb{F}$-vector space A which is also a ring, in such a way that the vector space structure and the ring structure satisfy the following compatibility condition for all $a, b \in A$ and $\lambda \in \mathbb{F}$:

$$\lambda(ab) = (\lambda a)b = a(\lambda b).$$

If $ab = ba$ for all $a, b \in A$ then A is called *commutative*.

Remark. We do not require rings to be unital. We call an algebra A unital if its underlying ring is unital, i.e. there exists a (necessarily unique) element $1_A \in A$ such that $a1_A = 1_A a = a$ for all $a \in A$.

Example 3.1.2. Let V be an $\mathbb{F}$-vector space and consider $\operatorname{End} V$, the set of all linear maps from V to V. Then $\operatorname{End} V$ is an $\mathbb{F}$-algebra where the multiplication is composition of maps. In particular, if $V = \mathbb{F}^n$ then $\operatorname{End} V \cong \operatorname{Mat}_n(\mathbb{F})$, the algebra of $n \times n$ matrices over $\mathbb{F}$.

Definition 3.1.3. A *normed algebra* is an algebra A which is normed as a vector space in such a way that the norm is *submultiplicative*, i.e.

$$\|ab\| \leq \|a\|\|b\| \quad \text{for all } a, b \in A.$$

If A is unital and nonzero, we require additionally that $\|1_A\| = 1$. We call A a *Banach algebra* if A is complete with respect to its norm.

Example 3.1.4. Let X be a normed space and consider $\mathcal{B}(X)$, the vector space of all bounded linear operators on X. This is a unital normed algebra where multiplication is given by composition of operators and the norm is the operator norm. If X is a Banach space then $\mathcal{B}(X)$ is a Banach algebra. As a special case, consider $X = \mathbb{F}$. Then $\mathcal{B}(X) \cong \mathbb{F}$ which is the most basic example of a Banach algebra.

Example 3.1.5. Let X be a normed space. The space $\mathcal{K}(X)$ of compact operators on X is a closed subspace of $\mathcal{B}(X)$. We know that the composition of two compact operators is compact, so $\mathcal{K}(X)$ is a normed algebra. If X is a Banach space then $\mathcal{K}(X)$ is a Banach algebra since it is a closed subspace of the Banach algebra $\mathcal{B}(X)$. However, $\mathcal{K}(X)$ is not unital unless X is finite-dimensional, in which case $\mathcal{K}(X) = \mathcal{B}(X)$.

Remark. A consequence of submultiplicativity is that

$$\|a^n\| = \|a \cdots a\| \le \|a\| \cdots \|a\| = \|a\|^n.$$

Example 3.1.6. Let $(X, \mathcal{A}, \mu)$ be a measure space. Then $A = L^\infty(X, \mathcal{A}, \mu)$ is a Banach algebra with respect to pointwise multiplication (this works both if we consider real-valued or complex-valued functions). We leave this as an exercise.

Example 3.1.7. Let G be a countable group and set $A = \ell^1(G)$ (either as a real or complex vector space). We define the *convolution* of $a, b \in A$ to be the function $a * b\colon G \to \mathbb{F}$ by

$$(a * b)(g) = \sum_{h \in G} a(h) b(h^{-1}g) \quad \text{for } g \in G.$$

We show that this is well-defined: Consider the function $F\colon G \times G \to \mathbb{F}$ given by $(g, h) \mapsto a(h)b(h^{-1}g)$. We claim that this function is in $\ell^1(G \times G)$. Indeed, by Tonelli's theorem we can compute $\sum_{(g,h) \in G \times G} |F(g, h)|$ as

$$\begin{aligned}
\sum_{g,h \in G} |F(g,h)| &= \sum_{h \in G} \sum_{g \in G} |a(h) b(h^{-1}g)| \\
&= \sum_{h \in G} \sum_{g \in G} |a(h)||b(g)| \\
&= \Big(\sum_{h \in G} |a(h)| \Big) \Big(\sum_{h \in G} |b(g)| \Big) \\
&= \|a\|_1 \|b\|_1 < \infty.
\end{aligned}$$

Since this is finite, we have by Fubini's theorem that the function $g \mapsto \sum_{h \in G} F(g, h)$ is everywhere defined, i.e. $(a * b)(g)$ is defined everywhere. Moreover, we can interchange the sums in the following computation

$$\|a * b\|_1 \le \sum_{g \in G} \sum_{h \in H} |a(h) b(h^{-1}g)| = \sum_{h \in G} \sum_{g \in H} |a(h) b(h^{-1}g)| = \|a\|_1 \|b\|_1.$$

This shows that the ℓ^1-norm is submultiplicative with respect to convolution. It is a straightforward exercise to check the rest of the algebra axioms. Thus, $\ell^1(G)$ is a Banach algebra with respect to convolution, and we will always view $\ell^1(G)$ as a Banach algebra in this way.

Remark. Let G be a *topological group*, that is, a group equipped with a topology in which the group multiplication and inversion are continuous maps. Let μ be a Radon measure on G which is left translation-invariant, i.e. $\mu(gS) = \mu(S)$ for all $g \in G$ and Borel sets S. Then one can define convolution on $L^1(G)$ via the formula

$$(a * b)(g) = \int_G a(h) b(h^{-1}g) \, \mathrm{d}\mu(h) \quad \text{for } a, b \in L^1(G).$$

One can then show that $L^1(G)$ becomes a Banach algebra. Haar's theorem states that such a measure μ always exists if the group is locally compact Hausdorff, and that the measure is unique up to a positive constant. However, that goes beyond the scope of this course.

Definition 3.1.8. Let A and B be algebras. A map $\phi\colon A \to B$ is called an *(algebra) homomorphism* if it is linear with respect to the underlying vector space structure of A and B and

$$\phi(aa') = \phi(a)\phi(a') \quad \text{for all } a, a' \in A.$$

If A and B are unital with multiplicative identities 1_A and 1_B respectively, then we call ϕ *unital* if in addition $\phi(1_A) = 1_B$. If A and B are normed algebras, we call ϕ *norm-decreasing* if

$$\|\phi(a)\| \leq \|a\| \quad \text{for all } a \in A.$$

Note that norm-decreasing homomorphisms of normed algebras are continuous.

Example 3.1.9. Let A be a nonzero unital algebra. Then the map $\mathbb{F} \to A$ given by $\lambda \mapsto \lambda 1_A$ is an injective unital algebra homomorphism. Thus $\mathbb{F}$ is embedded into A.

A subalgebra of an algebra A is a subspace B of A which is closed under multiplication: That is, if $a, b \in B$ then $ab \in B$. If A is unital with multiplicative identity 1_A, we call B a *unital* subalgebra if $1_A \in B$. Note that this is stronger than just requiring B to be unital as an algebra in itself.

Definition 3.1.10. Let A be an algebra. A subspace I of A is called a *left ideal* if $ab \in I$ whenever $a \in A$ and $b \in I$. It is called a *right ideal* if $ab \in I$ whenever $a \in I$ and $b \in A$. It is called a *(two-sided) ideal* if it is both left and right (often we will omit the word two-sided).

Recall that if X is a normed space and Y is a closed subspace, then the quotient space X/Y is a normed space with respect to the quotient norm

$$\|x + Y\| = \inf_{y \in Y} \|x + y\|.$$

If X is a Banach space then X/Y is a Banach space as well.

Proposition 3.1.11. *Let A be a normed algebra and let I be a closed ideal in A. Then the quotient space A/I becomes a normed algebra with respect to the quotient norm. Furthermore, if A is a Banach algebra then A/I is a Banach algebra.*

Proof. From the basics of ring theory we know that A/I is an algebra since I is an ideal in A. It remains to check that the quotient norm is submultiplicative. Let $a, b \in A$ and let $c, c' \in I$. Then $ac' + bc + cc' \in I$ since I is an ideal, so

$$\begin{aligned}
\|(a + I)(b + I)\| &= \|ab + I\| \\
&\leq \inf_{c,c' \in I} \|ab + ac + bc' + cc'\| \\
&= \inf_{c,c' \in I} \|(a + c)(b + c')\| \\
&\leq \inf_{c,c' \in I} \|a + c\|\|b + c'\| \\
&= \|a + I\|\|b + I\|.
\end{aligned}$$

This shows submultiplicativity. ∎

Example 3.1.12. Let X be a Banach space. The *Calkin algebra* of X is the Banach algebra $A = \mathcal{B}(X)/\mathcal{K}(X)$. Note that when X is finite-dimensional, then $A = \{0\}$.

3.2 Algebras of continuous functions

Throughout this section Ω denotes a locally compact Hausdorff topological space. This section concerns the $\mathbb{F}$-vector space $C(\Omega, \mathbb{F})$ of $\mathbb{F}$-valued, continuous functions on Ω, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. However, we shall eventually be interested in the complex-valued case, and will use $C(\Omega)$ to mean $C(\Omega, \mathbb{C})$ unless stated otherwise.

By $C_b(\Omega, \mathbb{F})$ we mean the subspace of $C(\Omega, \mathbb{F})$ consisting of bounded functions. Both $C(\Omega, \mathbb{F})$ and $C_b(\Omega, \mathbb{F})$ are algebras with respect to pointwise multiplication of functions. Furthermore, $C_b(\Omega, \mathbb{F})$ is complete with respect to the norm

$$\|f\|_\infty = \sup\{|f(t)| : t \in \Omega\} \quad \text{for } f \in C_b(\Omega),$$

and is thus a Banach algebra.

Definition 3.2.1. Let Ω be a locally compact Hausdorff space. A continuous function $f : \Omega \to \mathbb{F}$ is said to *vanish at infinity* if for every $\epsilon > 0$ there exists a compact set $K \subseteq \Omega$ such that $|f(t)| < \epsilon$ for all $t \in \Omega \backslash K$. We denote by $C_0(\Omega, \mathbb{F})$ the subset of $C(\Omega, \mathbb{F})$ consisting of functions that vanish at infinity.

Remark. Note that $C_0(\Omega, \mathbb{F}) \subseteq C_b(\Omega, \mathbb{F})$: Letting $\epsilon = 1$, we can find a compact set $K \subseteq \Omega$ such that $|f(t)| < 1$ for $t \in \Omega \setminus K$. Now f is continuous and K is compact, hence f is bounded on K, say by c. But then $|f(t)| \leq c'$ for all $t \in \Omega$, where $c' = \max\{1, c\}$.

Furthermore, when Ω is compact every continuous function vanishes at infinity trivially, so $C_0(\Omega, \mathbb{F}) = C(\Omega, \mathbb{F})$.

Example 3.2.2. Let $\Omega = \mathbb{R}$. A function $f \in C(\mathbb{R}, \mathbb{F})$ vanishes at infinity if and only if $\lim_{t\to\infty} |f(t)| = \lim_{t\to-\infty} |f(t)| = 0$.

Proposition 3.2.3. *Let Ω be a locally compact Hausdorff space. Then $C_0(\Omega, \mathbb{F})$ is a closed subalgebra of $C_b(\Omega, \mathbb{F})$, hence a Banach algebra (with respect to pointwise multiplication).*

Proof. We need to show that $C_0(\Omega, \mathbb{F})$ is a subspace of $C_b(\Omega, \mathbb{F})$. Let $f, g \in C_0(\Omega, \mathbb{F})$ and $\lambda, \mu \in \mathbb{F}$. Let $\epsilon > 0$. Assume first that $\lambda \neq 0$ and $\mu \neq 0$. Then there exist compact sets $K_1, K_2 \subseteq \Omega$ such that $|f(x)| < \epsilon/(2|\lambda|)$ for $x \in \Omega \setminus K_1$ and $|g(x)| < \epsilon/(2|\mu|)$ for $x \in \Omega \setminus K_2$. Thus, for $x \in \Omega \setminus (K_1 \cup K_2)$ we have that

$$|(\lambda f + \mu g)(x)| \leq |\lambda||f(x)| + |\mu||g(x)| < |\lambda| \cdot \frac{\epsilon}{2|\lambda|} + |\mu| \cdot \frac{\epsilon}{2|\mu|} = \epsilon.$$

If $\lambda = 0$ then it suffices to look at $x \in \Omega \setminus K_1$ and if $\mu = 0$ it suffices to look at $x \in \Omega \setminus K_2$. The proof that $C_0(\Omega, \mathbb{F})$ is closed under multiplication is similar, so we leave it as an exercise.

We now show that $C_0(\Omega, \mathbb{F})$ is closed in $C_b(\Omega, \mathbb{F})$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $C_0(\Omega, \mathbb{F})$ that converges in the ∞-norm to a function $f \in C_b(\Omega, \mathbb{F})$. Let $\epsilon > 0$. Then there exists $n \in \mathbb{N}$ such that $\|f - f_n\|_\infty < \epsilon/2$. Since f_n vanishes at infinity, we can find a compact set $K \subseteq \Omega$ such that $|f_n(t)| < \epsilon/2$ for $t \in \Omega \setminus K$. Hence, for all $t \in \Omega \setminus K$ we have that

$$|f(t)| \leq |f(t) - f_n(t)| + |f_n(t)| \leq \|f - f_n\|_\infty + |f_n(t)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence f vanishes at infinity, so $C_0(\Omega, \mathbb{F})$ is closed in $C_b(\Omega, \mathbb{F})$. This finishes the proof. ∎

Remark. The Banach algebra $C_0(\Omega, \mathbb{F})$ is unital if and only if Ω is compact. This will be left as an exercise.

We will now spend some time recalling the description of the dual space of $C_0(\Omega, \mathbb{F})$ in terms of Radon measures. This material is covered in MAT4410, so we omit proofs.

Let $\mathcal{B}$ denote the Borel σ-algebra of Ω. Recall that a *Radon measure* on Ω is a measure μ defined on $\mathcal{B}$ such that

(a) $\mu(U) = \sup\{\mu(K) : K \subseteq U \text{ compact}\}$ for all open sets U (inner regularity),

(b) $\mu(B) = \inf\{\mu(U) : B \subseteq U \text{ open }\}$ for all Borel sets B (outer regularity),

(c) $\mu(K) < \infty$ for all compact sets K.

We will only be concerned with *finite* Radon measures in this section, hence the last condition above is redundant.

Let μ be a real (also called signed) measure on the Borel σ-algebra $\mathcal{B}$ of Ω. Such a measure has a Jordan decomposition of the form $\mu = \mu_+ - \mu_-$ where μ_+ and μ_- are finite (nonnegative) measures on $\mathcal{B}$. The decomposition satisfies a certain uniqueness property. We say that μ is a *Radon measure* if both μ_+ and μ_- are Radon measures. Finally, if μ is a finite complex measure μ on the Borel σ-algebra of Ω then we can write $\mu = \operatorname{Re}\mu + i \operatorname{Im}\mu$ for finite real measures $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$. We then call μ a *Radon measure* if both $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$ are Radon measures. By an $\mathbb{F}$-valued measure we mean a real measure if $\mathbb{F} = \mathbb{R}$ and a complex measure if $\mathbb{F} = \mathbb{C}$.

The *total variation measure* of an $\mathbb{F}$-valued Radon measure μ on Ω is a (nonnegative) Radon measure on Ω defined by

$$|\mu|(S) = \sup\Big\{\sum_{j=1}^{\infty} |\mu(E_j)| : (E_j)_{j=1}^{\infty} \text{ is a countable partition of } S \text{ into Borel sets}\Big\}$$

for a Borel set $S \subseteq \Omega$. The *total variation* of μ is the number

$$\|\mu\| = |\mu|(\Omega).$$

Example 3.2.4. Let μ be a finite $\mathbb{F}$-valued Radon measure on Ω and let $g \in C_0(\Omega, \mathbb{F})$. Then the measure $g\mu$ on the Borel σ-algebra $\mathcal{B}$ of Ω given by

$$(g\mu)(S) = \int_S g \,\mathrm{d}\mu \quad \text{for } S \in \mathcal{B}$$

is a finite $\mathbb{F}$-valued Radon measure on Ω, and the total variation of $g\mu$ is given by $|g\mu| = |g||\mu|$.

Denote by $M(\Omega, \mathbb{F})$ the $\mathbb{F}$-vector space of $\mathbb{F}$-valued finite Radon measures on Ω. Then $M(\Omega, \mathbb{F})$ is a Banach space with respect to the total variation norm. Moreover, the map $\Phi\colon M(\Omega, \mathbb{F}) \to C_0(\Omega, \mathbb{F})^*$ given by

$$\Phi(\mu)(f) = \int_\Omega f \,\mathrm{d}\mu \quad \text{for } \mu \in M(\Omega, \mathbb{F}),\ f \in C_0(\Omega, \mathbb{F}),$$

is a surjective isometry (this is the Riesz–Markov–Kakutani theorem).

Definition 3.2.5. Let μ be a finite $\mathbb{F}$-valued Radon measure on Ω. The *support* of μ is the set $\operatorname{supp}(\mu)$ consisting of those $t \in \Omega$ such that $|\mu|(U) > 0$ for every open neighborhood U of t.

Example 3.2.6. (a) Note that if μ is a finite $\mathbb{F}$-valued Radon measure on Ω, then $\operatorname{supp}(\mu) = \Omega$ if and only if $|\mu|(U) > 0$ for all nonempty open sets $U \subseteq \Omega$. In particular, the Lebesgue measure on $[0, 1]$ has support equal to all of $[0, 1]$.

(b) Recall that given $t \in \Omega$, the *Dirac measure* δ_t is given by

$$\delta_t(S) = \begin{cases} 1 & \text{if } t \in S, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $\delta_t(U) = 1 > 0$ for all open neighborhoods U of t, so $t \in \operatorname{supp}(\delta_t)$. On the other hand, if $s \in \Omega$ and $s \neq t$, then since Ω is assumed to be Hausdorff we can find an open set U with $s \in U$ and $t \notin U$. Thus $\delta_t(U) = 0$ so $s \notin \operatorname{supp}(\delta_t)$. We conclude that $\operatorname{supp}(\delta_t) = \{t\}$.

Proposition 3.2.7. *Let μ be a finite $\mathbb{F}$-valued Radon measure on a locally compact Hausdorff space Ω, and let $t \in \Omega$. Then $\operatorname{supp}(\mu) = \{t\}$ if and only if $\mu = c\delta_t$ for some $c \in \mathbb{F} \setminus \{0\}$.*

Proof. As we saw in Example 3.2.6 (*b*), the Dirac measure δ_t has support equal to $\{t\}$. Consequently the support of $c\delta_t$ for $c \in \mathbb{F} \setminus \{0\}$ is also equal to $\{t\}$.

Conversely, suppose that $\operatorname{supp}(\mu) = \{t\}$. Let K be a compact set not containing t. Then for every $s \in K$, $s \notin \operatorname{supp}(\mu)$, so there exists an open neighborhood V_s of s such that $|\mu|(V_s) = 0$. But then the sets V_s for $s \in K$ cover K, so we can find finitely $s_1, \ldots, s_n \in K$ such that $V_{s_1}, \ldots, V_{s_n}$ cover K. It follows that

$$|\mu|(K) \leq \sum_{i=1}^{n} |\mu|(V_{s_i}) = 0.$$

By inner regularity of μ, it follows that $|\mu|(\Omega \setminus \{t\}) = 0$ since $\Omega \setminus \{t\}$ is an open set. Hence $|\mu|(B) = 0$ for all Borel sets B that do not contain t. This implies $|\mu(B)| = 0$ for these B (since we can partition B into just the single set U itself in the definition of the total variation measure), so $\mu(B) = 0$.

Assume now that B is a Borel set and $t \in B$. By what we already proved, $0 = \mu(B \setminus \{t\}) = \mu(B) - \mu(\{t\})$, so $\mu(B) = \mu(\{t\})$. Setting $c = \mu(\{t\})$, we have that $\mu(B) = c\delta_t(B)$ for all Borel sets B. We now have that $\mu = c\delta_t$. Since $|c| = |\mu|(\Omega) > 0$, it follows that $c \neq 0$. This finishes the proof. ■

Theorem 3.2.8 (The Stone–Weierstrass Theorem, real case)**.** *Let Ω be a compact Hausdorff space. Let A be a closed unital subalgebra of $C(\Omega, \mathbb{R})$ that separates points: That is, whenever $s, t \in \Omega$ and $s \neq t$, there exists $f \in A$ such that $f(s) \neq f(t)$. Then $A = C(\Omega, \mathbb{R})$.*

Before we proceed with the proof of Theorem 3.2.8, we need the following lemma:

Lemma 3.2.9. *Let μ be a real Radon measure on Ω and let $g \in C(\Omega, \mathbb{R})$ have the property that $g\mu = c\mu$ for some $c \in \mathbb{R}$. Then $g(t) = c$ for all $t \in \operatorname{supp}(\mu)$.*

Proof. Assume first that g and c are nonnegative. Suppose for a contradiction that $t \in \operatorname{supp}(\mu)$ and $g(t) \neq c$, say $g(t) > c$. By continuity of g we can find an open neighborhood of t and $n \in \mathbb{N}$ such that $g(t') > c + 1/n$ for all $t' \in U$. But then

$$c|\mu|(U) = |g\mu|(U) = \int_U g \, \mathrm{d}|\mu| \geq \int_U (c + 1/n) \, \mathrm{d}|\mu| \geq (c + 1/n)|\mu|(U).$$

Since $t \in \operatorname{supp}(\mu)$, we have that $|\mu|(U) > 0$, so the above equation gives $c \geq c + 1/n$, a contradiction. Hence $g(t) = c$ for all $t \in \operatorname{supp}(\mu)$. The proof in the case $g(t) < c$ is analogous.

If g is real-valued then $g\mu = c\mu$ implies that $|g||\mu| = |g\mu| = |c\mu| = |c||\mu|$. We can therefore apply the above proof to the nonnegative function $|g|$ and $|c| \geq 0$. ■

Proof of Theorem 3.2.8. We identify $C(\Omega, \mathbb{R})^*$ with $M(\Omega, \mathbb{R})$ using the Riesz–Markov–Kakutani Theorem. Consider the annihilator

$$A^\perp = \{\mu \in M(\Omega, \mathbb{R}) : \int_\Omega f \, \mathrm{d}\mu = 0 \text{ for all } f \in A\}.$$

Suppose for a contradiction that A is a proper subset of $C(\Omega, \mathbb{R})$.

Let $K = A^\perp \cap \bar{B}^*$, where $\bar{B}^*$ is the closed unit ball of $M(\Omega, \mathbb{R})$. In other words, K is the closed unit ball of the Banach space $A^\perp$. Thus, K is convex, and weak* compact by the Banach–Alaoglu Theorem (Theorem 2.5.7). By the Krein–Milman Theorem (strictly speaking by Lemma 2.6.6), K has an extreme point μ. It follows from an exercise $A^\perp$ is nonzero since A is proper, so by Example 2.6.3 we must have $\|\mu\| = 1$.

Since A is an algebra we have that $fg \in A$ whenever $f, g \in A$. Thus, for $g \in A$, we have

$$0 = \int_\Omega fg \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}(g\mu) \quad \text{for all } f \in A.$$

In other words $g\mu \in A^\perp$ whenever $g \in A$.

We claim that the support of μ consists of exactly one point. For a contradiction, suppose we can find $s, s' \in \operatorname{supp}(\mu)$ with $s \neq s'$. Since A separates points, we can find $h \in A$ such that $h(s) \neq h(s')$. Since A is unital, it contains the function constantly equal to 1. Scaling 1 by any scalar and using that A is closed under scalar multiplication, we have that all constant functions on Ω are in A. Using that A is closed under addition, we can add a large enough positive constant function to h to get a positive function which is still in A (remember that h is bounded). After dividing by another constant, we get a function $g \in A$ with $g(s) \neq g(s')$ and $0 < g(t) < 1$ for all $t \in \Omega$. Note also that $1 - g \in A$.

Set

$$c = \|g\mu\| = |g\mu|(\Omega) = \int_\Omega \mathrm{d}|g\mu| = \int_\Omega \mathrm{d}(|g||\mu|) = \int_\Omega g \, \mathrm{d}|\mu|.$$

Likewise, set $d = \|(1-g)\mu\| = \int_\Omega (1-g) \, \mathrm{d}|\mu|$. Since g and $1-g$ are both positive on Ω, c and d are positive numbers. Moreover,

$$c + d = \int_\Omega g \, \mathrm{d}|\mu| + \int_\Omega (1-g) \, \mathrm{d}\mu = 1.$$

As we have seen, both $g\mu$ and $(1-g)\mu$ are elements of $A^\perp$ since $g, (1-g) \in A$. Thus $g\mu/c, (1-g)\mu/d \in K$. Since μ is an extreme point of K, the equation

$$\mu = g\mu + (1-g)\mu = c\frac{g\mu}{c} + d\frac{(1-g)\mu}{d}$$

implies that $\mu = g\mu/c$. Using Lemma 3.2.9, we conclude that g is constant on $\operatorname{supp}(\mu)$. But then $g(s) = g(s')$ since $s, s' \in \operatorname{supp}(\mu)$, a contradiction. We conclude that $\operatorname{supp}(\mu) = t$ for some $t \in \Omega$.

By Proposition 3.2.7, $\mu = c\delta_t$ for some $c \in \mathbb{R} \setminus \{0\}$. Since $1 = \|\mu\| = |c|\|\delta_t\| = |c|$, we have that $c = \pm 1$. Possibly multiplying by -1, we can assume that $\mu = \delta_t$. Integrating the constant function $1 \in A$ with respect to $\delta_t \in A^\perp$, we obtain

$$1 = \delta_t(\Omega) = \int_\Omega 1 \, \mathrm{d}\delta_t = 0,$$

a contradiction. This finishes the proof. ■

Corollary 3.2.10 (The Stone–Weierstrass Theorem, complex case)**.** *Let Ω be a compact Hausdorff space. Let A be a closed unital subalgebra of $C(\Omega, \mathbb{C})$ that separates points and has the property that the complex conjugate $\overline{f}$ is in A for all $f \in A$. Then $A = C(\Omega, \mathbb{C})$.*

Proof. Let A be as in the statement of the corollary. $B = \{\operatorname{Re}(f) : f \in A\}$. Then $B \subsetneq C(\Omega, \mathbb{F})$ and B is closed in $C(\Omega, \mathbb{R})$. Moreover, $B \subseteq A$, since if $f \in A$ then $\operatorname{Re}(f) = (f + \overline{f})/2 \in A$. Finally, if $f, g \in A$ then $\operatorname{Re}(f) \cdot \operatorname{Re}(g)$ is a real-valued function in A, hence $\operatorname{Re}(f) \cdot \operatorname{Re}(g) = \operatorname{Re}(\operatorname{Re}(f) \cdot \operatorname{Re}(g) + i \cdot 0)$, so $\operatorname{Re}(f) \cdot \operatorname{Re}(g) \in B$. Similar arguments for linear combinations show that B is a subalgebra of $C(\Omega, \mathbb{R})$.

Now if $f \in A$ then $-if \in A$ so $\operatorname{Re}(-if) = \operatorname{Im}(f) \in B$. We show that B separates points: If $s, t \in \Omega$ and $s \neq t$ then we can find $f \in A$ such that $f(s) \neq f(t)$. But then either $\operatorname{Re}(f(s)) \neq \operatorname{Re}(f(t))$ or $\operatorname{Im}(f(s)) \neq \operatorname{Im}(f(t))$, which shows that B separates points. We can now apply the real version of the Stone–Weierstrass Theorem (Theorem 3.2.8) to conclude that $B = C(\Omega, \mathbb{R})$. Hence, if $f \in C(\Omega, \mathbb{C})$, then $\operatorname{Re}(f), \operatorname{Im}(f) \in B \subseteq A$, so

$$f = \operatorname{Re}(f) + i \operatorname{Im}(f) \in A.$$

■

We mention the following result for locally compact Hausdorff spaces, which we will not prove (although it is not so tedious to prove using the Stone–Weierstrass Theorem for compact spaces).

Proposition 3.2.11 (The Stone–Weierstrass Theorem, locally compact version)**.** *Let Ω be a locally compact Hausdorff space. Then the following hold:*

(a) Let A be a closed subalgebra of $C_0(\Omega, \mathbb{R})$ that separates points and vanishes nowhere*, i.e. for all $t \in \Omega$ there exists $f \in A$ such that $f(t) \neq 0$. Then $A = C_0(\Omega, \mathbb{R})$.*

(b) Let A be a closed subalgebra of $C_0(\Omega, \mathbb{C})$ that separates points, vanishes nowhere, and is closed under complex conjugation. Then $A = C_0(\Omega, \mathbb{C})$.

3.3 The spectrum

From now on, we assume that all algebras are complex.

Definition 3.3.1. Let A be a unital algebra. An element $a \in A$ is called *invertible* if there exists a (necessarily unique) $b \in A$ such that $ab = ba = 1_A$. We write $b = a^{-1}$ and denote by $GL(A)$ the set of all invertible elements of A.

If $a, b \in GL(A)$ then $(ab)(b^{-1}a^{-1}) = (b^{-1}a^{-1})(ab) = 1$, so $ab \in GL(A)$ with $(ab)^{-1} = b^{-1}a^{-1}$. Hence $GL(A)$ is a group under multiplication.

Example 3.3.2. (a) Consider $B(X)$ where X is a Banach space. Then $T \in \mathcal{B}(X)$ is invertible if and only if there exists another bounded linear operator S on X such that $ST = TS = I$. By the Open Mapping Theorem, we know that if T is a bounded bijection then its inverse is bounded, hence it suffices to check that T is a bijection to conclude that $T \in GL(\mathcal{B}(X))$.

(b) Consider $C(\Omega) = C(\Omega, \mathbb{C})$ where Ω is a compact Hausdorff space. If $f \in C(\Omega)$ then f is invertible if there exists $g \in C(\Omega)$ such that $fg = 1$, the function constantly equal to 1. In other words $f(t)g(t) = 1$ for all $t \in \Omega$. For this to be possible, it is necessary that f is nonvanishing, i.e. $f(t) \neq 0$ for all $t \in \Omega$. In that case we set $g(t) = 1/f(t)$ which is continuous and hence an element of $C(\Omega)$. We conclude that f is invertible if and only if it is nonvanishing.

Proposition 3.3.3 (Neumann series)**.** *Let A be a Banach algebra. If $a \in A$ with $\|a\| < 1$, then $1 - a$ is invertible, and the inverse is given by*

$$(1-a)^{-1} = \sum_{n=0}^{\infty} a^n.$$

Furthermore $\|(1-a)^{-1}\| \leq (1 - \|a\|)^{-1}$.

Proof. Note that since $\|a\| < 1$ we can use the geometric series formula as follows:

$$\sum_{n=0}^{\infty} \|a^n\| \leq \sum_{n=0}^{\infty} \|a\|^n = \frac{1}{1 - \|a\|}$$

Thus $b = \sum_{n=0}^{\infty} a^n$ is an absolutely convergent sequence so it converges by completeness of A. Now

$$ab = \sum_{n=0}^{\infty} aa^n = \sum_{n=0}^{\infty} a^{n+1} = \sum_{n=1}^{\infty} a^n = b - 1$$

and similarly $ba = b - 1$. Hence $(1-a)b = b(1-a) = 1$, so $1 - a \in GL(A)$, with $b = (1-a)^{-1}$. From the first estimate we also see that $\|b\| \leq (1 - \|a\|)^{-1}$. ■

As an application of Proposition 3.3.3 we have the following important proposition:

Proposition 3.3.4. *Let A be a Banach algebra. Then the following hold:*

(a) The set $GL(A)$ of invertible elements is an open subset of A.

(b) The inversion map $GL(A) \to GL(A)$ given by $a \mapsto a^{-1}$ is a homeomorphism.

Proof. (a): Let $a \in GL(A)$. We claim that the open ball centered at a with radius $\|a^{-1}\|^{-1}$ is contained in $GL(A)$. Indeed, suppose $\|a-b\| < \|a^{-1}\|^{-1}$. Then

$$\|1 - a^{-1}b\| = \|a^{-1}(a-b)\| \leq \|a^{-1}\|\|a-b\| < 1.$$

By Proposition 3.3.3 we have that $a^{-1}b = 1 - (1 - a^{-1}b) \in GL(A)$. But then $b = a(a^{-1}b) \in GL(A)$.

(b): We show continuity at $a \in GL(A)$. Let $\epsilon > 0$. If $b \in GL(A)$ satisfies $\|a-b\| < 1/(2\|a^{-1}\|)$ then

$$\|1 - a^{-1}b\| = \|a^{-1}(a-b)\| \leq \|a^{-1}\|\|a-b\| < \frac{1}{2}.$$

By Proposition 3.3.3 we have that $a^{-1}b$ is invertible, with

$$\|b^{-1}a\| = \|(a^{-1}b)^{-1}\| \leq \frac{1}{1 - \|1 - a^{-1}b\|} < 2.$$

Thus

$$\|b^{-1}\| = \|b^{-1}aa^{-1}\| \leq \|b^{-1}a\|\|a^{-1}\| < 2\|a^{-1}\|.$$

Now for any $b \in GL(A)$ we have that

$$\|a^{-1} - b^{-1}\| = \|a^{-1}(b-a)b^{-1}\| \leq \|a^{-1}\|\|b-a\|\|b^{-1}\|.$$

Hence, if $\|a-b\| < \delta := \min\{1/(2\|a^{-1}\|), \epsilon/(2\|a^{-1}\|^2)\}$, then $\|a^{-1} - b^{-1}\| < \epsilon$. ■

Definition 3.3.5. Let A be a unital algebra and let $a \in A$. The *spectrum* of a is the subset

$$\mathrm{sp}(a) = \{z \in \mathbb{C} : z1 - a \notin GL(A)\}$$

of the complex plane.

Example 3.3.6. (a) Consider $\mathcal{B}(X)$ for a normed space X. Then spectrum of an operator $T \in A$ is exactly the spectrum in the usual sense: That is, the set of $\lambda \in \mathbb{C}$ for which the operator $\lambda I - T$ is not invertible. In particular, if X is finite-dimensional then the spectrum of T is exactly the eigenvalues of T.

(b) Consider $C(\Omega) = C(\Omega, \mathbb{C})$, where Ω is a compact Hausdorff space. Let $f \in C(\Omega)$. Then $\lambda \in \mathrm{sp}(f)$ if and only if $\lambda - f$ is not invertible. From Example 3.3.2 (b) we know that this is equivalent to $\lambda - f$ having a zero, i.e. there exists $t \in \Omega$ such that $f(t) = \lambda$. But this is equivalent to $\lambda \in f(\Omega)$, the range of f. Thus $\mathrm{sp}(f) = f(\Omega)$.

Proposition 3.3.7. *Let A be a unital Banach algebra and let $a \in A$. Then* $\mathrm{sp}(a)$ *is a compact subset of* $\mathbb{C}$, *and* $|\lambda| \leq \|a\|$ *for all* $\lambda \in \mathrm{sp}(a)$.

Proof. We have that the maps $\iota\colon \mathbb{C} \to A$ and $T_a\colon A \to A$ given by $\iota(\lambda) = \lambda 1$ and $T_a(b) = b - a$, respectively, are continuous. Since $\mathbb{C} \setminus \mathrm{sp}(a) = \iota^{-1}(T_a^{-1}(GL(A)))$ and $GL(A)$ is open by Proposition 3.3.4 (a), it follows that $\mathbb{C} \setminus \mathrm{sp}(a)$ is open. Hence $\mathrm{sp}(a)$ is closed.

Furthermore, if $|z| > \|a\|$ then $\|z^{-1}a\| < 1$ so $1 - z^{-1}a \in GL(A)$ by Proposition 3.3.3. But then $z1 - a = z(1 - z^{-1}a)$ is also in $GL(A)$, so $\lambda \notin \mathrm{sp}(a)$. This shows that $|\lambda| \leq \|a\|$ for all $\lambda \in \mathrm{sp}(a)$. In particular, $\mathrm{sp}(a)$ is bounded, and since it is also closed it must be compact by the Heine–Borel Theorem. ■

Consider for a moment the Banach algebra $\mathcal{B}(X)$ of linear maps on an n-dimensional normed space X. After identifying X with $\mathbb{C}^n$, we have that $\mathcal{B}(X) \cong \mathrm{Mat}_n(\mathbb{C})$, the algebra of complex $n \times n$ matrices. In this setting we know that the spectrum of a matrix M consists exactly of its eigenvalues, which are the roots of its characteristic polynomial. By the fundamental theorem of algebra we know that this polynomial has a root; hence M has an eigenvalue. We will prove a vast generalization of this theorem to the setting of Banach algebras. First we need the following lemma.

Lemma 3.3.8. *Let A be a unital Banach algebra, let $a \in A$ and let $\phi \in A^*$. Then the complex-valued function*

$$f(z) = \phi((z1 - a)^{-1})$$

is holomorphic on the open set $\mathbb{C} \setminus \mathrm{sp}(a)$.

Proof. Note first that $\mathbb{C} \setminus \mathrm{sp}(a)$ is in fact open, since $\mathrm{sp}(a)$ is closed by Proposition 3.3.7.

Let $z, z_0 \in \mathbb{C} \setminus \mathrm{sp}(a)$ with $z \neq z_0$. We then have that

$$\begin{aligned} \frac{f(z) - f(z_0)}{z - z_0} &= \frac{\phi((z1-a)^{-1} - (z_0 1 - a)^{-1})}{z - z_0} \\ &= \frac{\phi((z1-a)^{-1}((z_0 1 - a) - (z1 - a))(z_0 1 - a)^{-1})}{z - z_0} \\ &= -\phi((z1-a)^{-1}(z_0 1 - a)^{-1}). \end{aligned}$$

Letting $z \to z_0$ and using the continuity of ϕ we see that $f'(z_0)$ exists, with $f'(z_0) = -\phi((a1 - z_0)^{-2})$. This shows that f is holomorphic at z_0. ■

Theorem 3.3.9. *Let A be a unital Banach algebra and let $a \in A$. Then* $\mathrm{sp}(a)$ *is nonempty.*

Proof. Suppose for a contradiction that $\mathrm{sp}(a)$ is empty. Let ϕ be a bounded linear functional on A. Then the function $f(z) = \phi((z1 - a)^{-1})$ for $z \in \mathbb{C}$ is holomorphic on all of $\mathbb{C}$ by Lemma 3.3.8.

If $|z| > \|a\|$ then we know from Proposition 3.3.7 that $z \notin \mathrm{sp}(a)$. Using the estimate from Proposition 3.3.3, we have that

$$\|(z1-a)^{-1}\| = |z|^{-1}\|(1 - z^{-1}a)^{-1}\| \leq \frac{1}{|z|(1 - \|z^{-1}a\|)} = \frac{1}{|z| - \|a\|}. \tag{3.1}$$

Thus, when $|z| > \|a\|$ we have that

$$|f(z)| = |\phi((z-a)^{-1})| \leq \|\phi\| \|(z-a)^{-1}\| \leq \frac{\|\phi\|}{|z| - \|a\|}. \tag{3.2}$$

Since f is continuous, it is bounded on the compact set $\{z \in \mathbb{C} : |z| \leq \|a\| + 1\}$. On the other hand, (3.2) shows that $|f(z)| \leq \|\phi\|$ for z in the set $\{z \in \mathbb{C} : |z| > \|a\| + 1\}$ as well. Thus, f is holomorphic and bounded on all of $\mathbb{C}$. By Liouville's theorem from complex analysis it follows that f is constant. Hence $\phi((z-a)^{-1}) = \phi((w-a)^{-1})$ for all $z, w \in \mathbb{C}$. Since $\phi \in A^*$ was arbitrary, it follows from a corollary of the Hahn–Banach Theorem that $(z-a)^{-1} = (w-a)^{-1}$ for all $z, w \in \mathbb{C}$. But this leads to a contradiction, as we can set e.g. $z = 0$ and $w = 1$ to obtain $-a = 1 - a$ or $0 = 1$. ■

Corollary 3.3.10 (Gelfand–Mazur Theorem)**.** *Let A be a unital Banach algebra in which very nonzero element is invertible. Then $A \cong \mathbb{C}$.*

Proof. Let $a \in A$. If $\lambda \in \operatorname{sp}(a)$ then $a - \lambda 1_A$ is not invertible. But then it must be zero, so $a = \lambda 1_A$. Thus A consists only of complex scalar multiples of the multiplicative identity 1_A, so the isometric homomorphism $\mathbb{C} \to A$ given by $\lambda \mapsto \lambda 1_A$ is surjective. It follows that $A \cong \mathbb{C}$. ■

Let a be an element of an algebra A and let $p(z)$ be a complex polynomial in the variable z, say $p(z) = \sum_{k=0}^{n} \lambda_k z^k$ for $\lambda_0, \ldots, \lambda_k \in \mathbb{C}$. We then set

$$p(a) := \sum_{k=0}^{n} \lambda_k a^k.$$

Proposition 3.3.11 (Spectral Mapping Theorem for polynomials)**.** *Let A be a normed, unital algebra and let $a \in A$. Let $p(z)$ be a complex polynomial in the variable z. Then*

$$\operatorname{sp}(p(a)) = p(\operatorname{sp}(a)) := \{p(\lambda) : \lambda \in \operatorname{sp}(a)\}.$$

Proof. If $p(z)$ is the zero polynomial then $\operatorname{sp}(p(a)) = \operatorname{sp}(0) = \{\lambda \in \mathbb{C} : \lambda 1 - 0 \notin GL(A)\} = \{0\} = p(\operatorname{sp}(a))$, so assume that $p(z)$ is nonzero.

First we show that $p(\operatorname{sp}(a)) \subseteq \operatorname{sp}(p(a))$. Let $\lambda \in \mathbb{C}$ and suppose that $p(\lambda) \notin \operatorname{sp}(p(a))$. Then $p(a) - p(\lambda)1$ has an inverse b. Since λ is a root of the polynomial $p(z) - p(\lambda)$, there exists a polynomial $q(z)$ such that $p(z) - p(\lambda) = q(z)(z - \lambda)$. But then $p(a) - p(\lambda)1 = q(a)(a - \lambda 1)$, so $bq(a)$ is an inverse of $a - \lambda 1$. Hence $\lambda \notin \operatorname{sp}(a)$.

We now show that $\operatorname{sp}(p(a)) \subseteq p(\operatorname{sp}(a))$. Let $\mu \in \operatorname{sp}(p(a))$. By the Fundamental Theorem of Algebra, we can write the polynomial $p(z) - \mu$ as a product of linear factors

$$p(z) - \mu = c(z - \lambda_1) \cdots (z - \lambda_n)$$

where $c, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Since $p(z)$ is assumed nonzero, c must be nonzero. But then we have

$$p(a) - \mu 1 = c(a - \lambda_1 1) \cdots (a - \lambda_n 1).$$

Since $p(a) - \mu 1$ is not invertible, at least one of the terms $a - \lambda_k 1$ cannot be invertible. But then $\lambda_k \in \operatorname{sp}(a)$, and

$$p(\lambda_k) = (p(\lambda_k) - \mu) + \mu = \mu,$$

which shows that $\mu \in p(\operatorname{sp}(a))$. This finishes the proof. ■

Definition 3.3.12. Let A be a unital Banach algebra, and let $a \in A$. Then the *spectral radius* of a is the number

$$r(a) = \sup\{|\lambda| : \lambda \in \operatorname{sp}(a)\}.$$

Remark. By Theorem 3.3.9 and Proposition 3.3.7, the spectrum is compact and nonempty, so the supremum in the definition of $r(a)$ is well-defined. It also follows from Proposition 3.3.7 that

$$r(a) \leq \|a\| \quad \text{for all } a \in A.$$

Theorem 3.3.13 (The Spectral Radius Formula)**.** *Let A be a unital Banach algebra and let $a \in A$. Then*

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n}.$$

Proof. Let $a \in A$. We will show that

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq r(a) \leq \liminf_{n\to\infty} \|a^n\|^{1/n}.$$

By Proposition 3.3.11 we have that $\lambda^n \in \operatorname{sp}(a^n)$ for all $\lambda \in \operatorname{sp}(a)$ and $n \in \mathbb{N}$. Hence, using Proposition 3.3.7, we have that $|\lambda|^n = |\lambda^n| \leq \|a^n\|$ whenever $\lambda \in \operatorname{sp}(a)$. This gives $|\lambda| \leq \liminf_{n\to\infty} \|a^n\|^{1/n}$, and thus

$$r(a) \leq \liminf_{n\to\infty} \|a^n\|^{1/n}.$$

We will now show the other inequality. Let $\phi \in A^*$ and consider the function $f(z) = \phi((z1 - a)^{-1})$, which is holomorphic on $\mathbb{C} \setminus \operatorname{sp}(a)$ by Lemma 3.3.8. Consider the annulus $S = \{z \in \mathbb{C} : r(a) < |z|\}$. Since $S \subseteq \mathbb{C} \setminus \operatorname{sp}(a)$, f is holomorphic on S.

Now if z belongs to the annulus $S' = \{z \in \mathbb{C} : \|a\| < |z|\}$, so that $\|z^{-1}a\| < 1$, then by Proposition 3.3.3 we have that

$$(z1 - a)^{-1} = \frac{1}{z}(1 - z^{-1}a)^{-1} = \frac{1}{z}\sum_{n=0}^{\infty}(z^{-1}a)^n = \sum_{n=0}^{\infty}\frac{1}{z^{n+1}}a^n.$$

It follows that

$$f(z) = \phi\Big(\sum_{n=0}^{\infty}\frac{1}{z^{n+1}}a^n\Big) = \sum_{n=0}^{\infty}\frac{\phi(a^n)}{z^{n+1}} \tag{3.3}$$

for $z \in S'$. The annulus S' is a subset of the annulus S (by Proposition 3.3.7), and f has a Laurent series on the bigger annulus S since it is holomorphic there. By uniqueness of Laurent series, the Laurent series (3.3) holds on the bigger annulus S. By the absolute convergence of a Laurent series, we then have that

$$\sum_{n=0}^{\infty}\Big|\frac{\phi(a^n)}{z^{n+1}}\Big| < \infty$$

whenever $|z| > r(a)$, so the sequence $(|\phi(a^n)/z^{n+1}|)_{n=0}^{\infty}$ is bounded for all $z \in \mathbb{C}$ with $|z| > r(a)$. This holds for all $\phi \in A^*$.

Consider now the linear isometry $\Phi\colon A \to A^{**}$ of A into its bidual. We have shown that

$$\sup_{n\geq 0} |\Phi(a^n/z^{n+1})(\phi)| < \infty \quad \text{for all } \phi \in A^* \text{ and } z \in \mathbb{C} \text{ with } |z| > r(a).$$

By the Uniform Boundedness Principle, we conclude that for every $z \in \mathbb{C}$ with $|z| > r(a)$, there exists $M_z > 0$ such that

$$\|a^n/z^{n+1}\| = \|\Phi(a^n/z^{n+1})\| \leq M_z \quad \text{for all } n \geq 0.$$

But then $\|a^n\| \leq M_z|z|^{n+1}$, so $\|a^n\|^{1/n} \leq M_z^{1/n}|z|^{1+1/n}$. Taking the limit superior, we obtain

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq \limsup_{n\to\infty} M_z^{1/n}|z|^{1+1/n} = \lim_{n\to\infty} M_z^{1/n}|z|^{1+1/n} = |z|,$$

which holds provided $|z| > r(a)$. Taking the infimum over all $z \in \mathbb{C}$ with $|z| > r(a)$ in the above inequality, we obtain

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq r(a).$$

This finishes the proof. ■

3.4 The Gelfand transform

Definition 3.4.1. Let A be a unital Banach algebra. A proper ideal I in A is called *maximal* if it is maximal among all the proper ideals of A ordered with respect to inclusion, i.e. if whenever J is a proper ideal of A and $I \subseteq J$, then $I = J$.

Proposition 3.4.2. *Let A be a unital Banach algebra. Then the following hold:*

(a) If I is an ideal of A, then $\mathrm{Cl}(I)$ is an ideal of A. If I is a proper ideal of A, then $\mathrm{Cl}(I)$ is a proper ideal of A.

(b) If I is a maximal ideal of A, then I is closed.

(c) If A is commutative, then every proper ideal is contained in a maximal ideal.

Proof. (a): Suppose I is an ideal of A. Then $\mathrm{Cl}(I)$ is a linear subspace, being the closure of a linear subspace. Let $a \in \mathrm{Cl}(I)$. Then we can find a sequence $(a_n)_{n\in\mathbb{N}}$ in I that converges to a. Let $b \in A$. Then $ab = \lim_{n\to\infty} a_n b$. Since $a_n b \in I$ for each $n \in \mathbb{N}$, it follows that $ab \in \mathrm{Cl}(I)$. This shows that $\mathrm{Cl}(I)$ is an ideal of A. Suppose that $\mathrm{Cl}(I) = A$. Then $1 \in \mathrm{Cl}(I)$, so we can find $a \in I$ with $\|1 - a\| < 1$. By Proposition 3.3.3 we have that $a \in GL(A)$. But then $a^{-1}a = 1 \in I$, so $b = b1 \in I$ for all $b \in A$. Thus $I = A$. This shows that if I is proper, then $\mathrm{Cl}(I)$ must be proper as well.

(b): Suppose I is a maximal ideal. From (a) we know that $\mathrm{Cl}(I)$ is an ideal of A as well. Since $I \subseteq \mathrm{Cl}\, I$, it follows from maximality that $I = \mathrm{Cl}(I)$, so I must be closed.

(c): This is a common application of Zorn's lemma which is commonly taught in a first course in abstract algebra, so we only outline the proof here: Let I be an ideal of A and order the set of all proper ideals that contain I with respect to inclusion. Show that every chain has an upper bound, and apply Zorn's lemma. ■

Example 3.4.3. Consider the commutative, unital Banach algebra $C(\Omega)$ for a compact Hausdorff space Ω. For a closed S of Ω, set

$$I(S) = \{f \in C(\Omega) : f(t) = 0 \text{ for all } t \in S\}.$$

It will be given as an exercise to show that the following hold:

(a) $I(S)$ is a closed ideal of $C(\Omega)$,

(b) If S and S' are closed subsets of Ω, then $S \subseteq S'$ if and only if $I(S) \supseteq I(S')$,

(c) Every closed ideal of $C(\Omega)$ is of the form $I(S)$ for some closed subset S in Ω.

Now a proper, closed ideal I of $C(\Omega)$ is maximal if and only if whenever $I \subseteq J$ for a proper ideal J (which we can assume is closed), then $J = I$. Writing $I = I(S)$ for a closed subset S of Ω and using the above, I is maximal if and only if S is a nonempty closed subset of Ω, and whenever $S' \subseteq S$ for a nonempty closed subset S' of S, then $S = S'$. This happens if and only if $S = \{t\}$ for some $t \in \Omega$. We conclude that the maximal ideals of $C(\Omega)$ are exactly the sets

$$I(\{t\}) = \{f \in C(\Omega) : f(t) = 0\} \quad \text{for } t \in \Omega.$$

Definition 3.4.4. Let A be a unital normed algebra. A unital algebra homomorphism from A to $\mathbb{C}$ is called a *character* on A. We denote by $\Delta(A)$ the set of all characters on A.

Remark. Requiring that an algebra homomorphism $\tau\colon A \to \mathbb{C}$ is unital is equivalent to requiring that it is nonzero: If τ is not identically zero then $\tau(a) \neq 0$ for some $a \in A$, so $0 \neq \tau(a) = \tau(a1) = \tau(a)\tau(1)$ which forces $\tau(1) = 1$. Requiring that τ is unital is also equivalent to surjectivity, since if $\tau(1) = 1$ then $\lambda = \lambda\tau(1) = \tau(\lambda 1)$ for all $\lambda \in \mathbb{C}$. Conversely, if τ is surjective then it is nonzero.

Example 3.4.5. Consider $C(\Omega)$ for a compact Hausdorff space Ω. Given $t \in \Omega$, we define a function $\mathrm{ev}_t\colon A \to \mathbb{C}$ by $\mathrm{ev}_t(f) = f(t)$ for $f \in C(\Omega)$. We know already that ev_t is a bounded linear functional on $C(\Omega)$ (it corresponds to the dirac measure δ_t). However, ev_t is also multiplicative, as

$$\mathrm{ev}_t(fg) = (fg)(t) = f(t)g(t) = \mathrm{ev}_t(f)\,\mathrm{ev}_t(g) \quad \text{for all } f, g \in C(\Omega).$$

We also have that $\mathrm{ev}_t(1) = 1(t) = 1$, so ev_t is a character on $C(\Omega)$.

Proposition 3.4.6. *Let A be a nonzero unital Banach algebra. Then $\Delta(A)$ is a weak* closed subset of the closed unit ball of A^*, hence compact in the weak* topology.*

Proof. We claim that if $a \in A$ and $\tau \in \Delta(A)$, then $\tau(a) \in \mathrm{sp}(a)$. Indeed, since τ is a unital algebra homomorphism, it maps invertible elements to invertible elements. But $\tau(\tau(a)1 - a) = \tau(a)\tau(1) - \tau(a) = 0$ which is not invertible, hence $\tau(a)1 - a \notin GL(A)$. Thus $\tau(a) \in \mathrm{sp}(a)$.

It follows from Proposition 3.3.7 that $|\tau(a)| \leq \|a\|$ for all $a \in A$, so τ is bounded, with $\|\tau\| \leq 1$. Since $|\tau(1)| = 1$ and $\|1\| = 1$, it follows that $\|\tau\| = 1$, so we have shown that $\Delta(A)$ is a subset of the closed unit ball of A^*.

Finally, suppose $(\tau_\lambda)_{\lambda \in \Lambda}$ is a net in $\Delta(A)$ that converges weak* to ϕ in the closed unit ball of A^*. If $a, b \in A$, then

$$\phi(ab) = \lim_\lambda \tau_\lambda(ab) = \lim_\lambda \tau_\lambda(a)\tau_\lambda(b) = (\lim_\lambda \tau_\lambda(a))(\lim_\lambda \tau_\lambda(b)) = \phi(a)\phi(b).$$

Furthermore, $\phi(1) = \lim_\lambda \tau_\lambda(1) = \lim_\lambda 1 = 1$. This shows that $\phi \in \Delta(A)$, so $\Delta(A)$ is a weak* closed subset of the closed unit ball of A^*. The weak* compactness of $\Delta(A)$ now follows from the Banach–Alaoglu Theorem (Theorem 2.5.7). ■

From now on, we will consider $\Delta(A)$ as a topological space with the weak* topology. By Proposition 3.4.6, $\Delta(A)$ is a compact Hausdorff space. We call $\Delta(A)$ the *character space* of A.

Lemma 3.4.7. *Let A be a nonzero unital commutative Banach algebra, and let I be a maximal ideal of A. Then A/I is isomorphic to $\mathbb{C}$.*

Proof. Pick $a_0 \in A \setminus I$ and set

$$J = \{aa_0 + b : a \in A, b \in I\}.$$

We claim that J is an ideal of A: For $a, a' \in A$, $b, b' \in I$ and $\lambda, \mu \in \mathbb{C}$ we have that

$$\lambda(aa_0 + b) + \mu(a'a_0 + b') = (\lambda a + \mu a')a_0 + (\lambda b + \mu b') \in J.$$

Furthermore, if $c \in A$ then

$$c(aa_0 + b) = (ca)a_0 + (cb)$$

which is in J since $cb \in I$. We also have that $I \subseteq J$, since elements of the form $0 \cdot a_0 + b = b$ are in J for $b \in I$. However, $a_0 \notin I$ but $a_0 = 1a_0 + 0 \in J$. It follows that $I \subsetneq J$, so by maximality of J we conclude that $J = A$. We can then find $a \in A$ and $b \in I$ such that $1 = aa_0 + b$. Thus $1 + I = aa_0 + I = (a + I)(a_0 + I)$, so $a_0 + I$ is invertible in A/I. We have thus shown that every nonzero element in A/I is invertible, so $A/I \cong \mathbb{C}$ by the Gelfand–Mazur Theorem (Corollary 3.3.10). ■

Proposition 3.4.8. *Let A be a nonzero, commutative, unital Banach algebra. Then the following hold:*

(a) The map $\tau \mapsto \mathrm{Ker}(\tau)$ is a bijection between $\Delta(A)$ and the set of maximal ideals of A.

(b) An element $a \in A$ is invertible if and only if $\tau(a) \neq 0$ for all $\tau \in \Delta(A)$, and

$$\mathrm{sp}(a) = \{\tau(a) : \tau \in \Delta(A)\} \quad \text{for all } a \in A.$$

Proof. (a): Let $\tau \in \Delta(A)$. We must show that $\mathrm{Ker}(\tau)$ is a maximal ideal of A. Since $1 \notin \mathrm{Ker}\,\tau$, we know that $\mathrm{Ker}(\tau)$ is a proper ideal of A. Let $\mathrm{Ker}(\tau) \subseteq I$ for an ideal I of A. Assuming that $\mathrm{Ker}(\tau) \neq I$, we can let $a \in I \setminus \mathrm{Ker}(\tau)$. Then $\tau(a) \neq 0$, while $\tau(1 - \tau(a)^{-1}a) = 1 - \tau(a)^{-1}\tau(a) = 0$. Thus $1 - \tau(a)^{-1}a \in \mathrm{Ker}(\tau) \subseteq I$, so $1 = (1 - \tau(a)^{-1}a) + \tau(a)^{-1}a \in I$. This shows that $I = A$, so $\mathrm{Ker}(\tau)$ must be maximal.

We show injectivity of the map: Suppose that $\tau_1, \tau_2 \in \Delta(A)$ satisfy $\mathrm{Ker}(\tau_1) = \mathrm{Ker}(\tau_2)$. Let $a \in A$. Then $a - \tau_1(a)1 \in \mathrm{Ker}(\tau_1)$, so by assumption $a - \tau_1(a)1 \in \mathrm{Ker}(\tau_2)$. But then

$$0 = \tau_2(a - \tau_1(a)1) = \tau_2(a) - \tau_1(a)\tau_2(1) = \tau_2(a) - \tau_1(a),$$

so $\tau_1(a) = \tau_2(a)$. Since $a \in A$ was arbitrary, we conclude that $\tau_1 = \tau_2$.

Finally, we show surjectivity: Let I be a maximal ideal of A. By Lemma 3.4.7, the quotient A/I is isomorphic to $\mathbb{C}$. Moreover, the quotient map $A \to A/I$ is a unital algebra homomorphism. We then have a unital algebra homomorphism $\tau \colon A \to A/I \cong \mathbb{C}$, i.e. a character on A, and $\mathrm{Ker}(\tau) = I$. This shows that $\Delta(A)$ is in bijection with the maximal ideals of A.

(b): If a is invertible and $\tau \in \Delta(A)$, then $\tau(a)$ must be invertible in $\mathbb{C}$, so $\tau(a) \neq 0$. Conversely, suppose a is not invertible. Consider the set $I = \{ab : b \in A\}$, which is an ideal since A is commutative. Furthermore, $a = a1 \in I$, and I is proper since $1 \notin I$ (otherwise we would have $1 = ab$ for some $b \in A$, which contradicts that a is not invertible). By Proposition 3.4.2, I is contained in a maximal ideal J. By (a), there exists $\tau \in \Delta(A)$ such that $\mathrm{Ker}(\tau) = J$. Since $a \in I \subseteq J$, we have that $\tau(a) = 0$.

We already saw that $\tau(a) \in \mathrm{sp}(a)$ for all $a \in A$, since $\tau(a)1 - a \in \mathrm{Ker}(\tau)$. Conversely, let $\lambda \in \mathrm{sp}(a)$. Then $\lambda 1 - a$ is not invertible, so by (b) there exists $\tau \in \Delta(A)$ with $\tau(\lambda 1 - a) = 0$. But then $\tau(a) = \lambda$, which finishes the proof. ■

We can now determine the character space of the Banach algebra $C(\Omega)$: It is simply the space Ω itself.

Proposition 3.4.9. *Let Ω be a compact Hausdorff space. Define a map $\Psi\colon \Omega \to \Delta(C(\Omega))$ by*

$$\Psi(t) = \mathrm{ev}_t \quad \text{for } t \in \Omega.$$

Then Ψ is a homeomorphism.

Proof. We begin by showing injectivity of Ψ: Let $s, t \in \Omega$ and suppose that $s \neq t$. Since Ω is compact Hausdorff, it is normal. By Urysohn's lemma, we can find a continuous function $f\colon \Omega \to \mathbb{C}$ such that $f(s) = 1$ and $f(t) = 0$. But then $\Psi(s)(f) = f(s) \neq f(t) = \Psi(t)(f)$, so $\Psi(s) \neq \Psi(t)$.

Next, we show surjectivity of Ψ: Let $\tau \in \Delta(C(A))$. Then $\mathrm{Ker}(\tau)$ is a maximal ideal by Proposition 3.4.8 (a), so by Example 3.4.3 we must have $\mathrm{Ker}(\tau) = I(\{t\})$ for some $t \in \Omega$. But $I(\{t\}) = \mathrm{Ker}(\mathrm{ev}_t)$, so again by Proposition 3.4.8 we conclude that $\tau = \mathrm{ev}_t$.

The proof of continuity of Ψ goes as follows: If $(t_\lambda)_{\lambda \in \Lambda}$ is a net in Ω that converges to $t \in \Omega$, then $\Psi(t_\lambda)(f) = f(t_\lambda) \to f(t) = \Psi(t)(f)$ for all $f \in C(\Omega)$. Hence $\Psi(t_\lambda) \to \Psi(t)$ in the weak* topology.

Now since Ψ is a continuous bijection from a compact space to a Hausdorff space, it follows from general topology that Ψ is a homeomorphism.

■

As before, let A be a commutative unital Banach algebra. Given $a \in A$, we define $\widehat{a}\colon \Delta(A) \to \mathbb{C}$ by

$$\widehat{a}(\tau) = \tau(a) \quad \text{for all } \tau \in \Delta(A).$$

Note that $\widehat{a}$ is continuous: Indeed, if $(\tau_\lambda)_{\lambda \in \Lambda}$ is a net in $\Delta(A)$ that converges weak* to $a \in \Delta(A)$, then $\widehat{a}(\tau_\lambda) = \tau_\lambda(a) \to \tau(a) = \widehat{a}(\tau)$. This shows that $\widehat{a} \in C(\Delta(A))$.

Definition 3.4.10. Let A be a commutative, unital Banach algebra. We define the *Gelfand transform* of A to be the map

$$\Gamma\colon A \to C(\Delta(A))$$

given by

$$\Gamma(a) = \widehat{a} \quad \text{for } a \in A.$$

Theorem 3.4.11. *Let A be a commutative, unital Banach algebra. Then the following hold:*

(a) The Gelfand transform of A is a unital algebra homomorphism.

(b) For all $a \in A$ we have that

$$\mathrm{sp}(\Gamma(a)) = \mathrm{sp}(a).$$

In particular, $\|\Gamma(a)\|_\infty = r(a)$, so that Γ is norm-decreasing, with $\|\Gamma\| = 1$.

Proof. (a): It is straightforward to check that Γ is a unital algebra homomorphism: Let $a, b \in A$ and $\lambda, \mu \in \mathbb{C}$. For $\tau \in \Delta(A)$ we have that

$$\Gamma(\lambda a + \mu b)(\tau) = \tau(\lambda a + \mu b) = \lambda\tau(a) + \mu\tau(b) = \lambda\Gamma(a)(\tau) + \mu\Gamma(b)(\tau) = \Gamma(\lambda a + \mu b)(\tau).$$

This shows that Γ is linear. Furthermore,

$$\Gamma(ab)(\tau) = \tau(ab) = \tau(a)\tau(b) = \Gamma(a)(\tau)\Gamma(b)(\tau)$$

shows that Γ preserves multiplication. Finally, $\Gamma(1)(\tau) = \tau(1) = 1$, so $\Gamma(1)$ is the function on $\Delta(A)$ constantly equal to 1. This shows that Γ is unital.

(b): Using Proposition 3.4.8 (b) we have that

$$\mathrm{sp}(\Gamma(a)) = \{\Gamma(a)(\tau) : \tau \in \Delta(A)\} = \{\tau(a) : \tau \in \Delta(A)\} = \mathrm{sp}(a).$$

Consequently $\|\Gamma(a)\|_\infty = r(\Gamma(a)) = r(a)$. Since $r(a) \leq \|a\|$, this shows that Γ is norm-decreasing, so $\|\Gamma\| \leq 1$. Since Γ is unital, it follows that $\|\Gamma\| = 1$. ■

Remark. The Gelfand transform is neither injective or surjective in general.

3.5 Convolution algebras of abelian groups

Let G be a countable group. Recall that $\ell^1(G)$ is a unital Banach algebra with respect to convolution, and that $\ell^1(G)$ is commutative if and only if G is abelian. Thus, if G is abelian, we might ask what the character space $\Delta(\ell^1(G))$ looks like, and what the Gelfand transform $\Gamma\colon \ell^1(G) \to C(\Delta(\ell^1(G)))$ does.

We denote by $\mathbb{T}$ the *circle group*, which can be thought of as the multiplicative subgroup

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$$

of the complex numbers (with the subspace topology), or the quotient group $\mathbb{R}/\mathbb{Z}$ (with the quotient topology). The isomorphism between the two comes from the map $\mathbb{R} \to \mathbb{C}$ given by $t \mapsto e^{2\pi i t}$, which has image equal to $\{z \in \mathbb{C} : |z| = 1\}$ and kernel equal to $\mathbb{Z}$. Note that $\mathbb{T}$ is compact, being a closed and bounded subset of $\mathbb{C}$.

Definition 3.5.1. Let G be an abelian countable group. A group homomorphism $G \to \mathbb{T}$ is called a *character* on G. The set $\widehat{G}$ of all characters of G is called the *Pontryagin dual* of G.

The Pontryagin dual $\widehat{G}$ of an abelian group G is a group itself with respect to pointwise multiplication. The inverse of a character $\omega\colon G \to \mathbb{T}$ is given by the pointwise conjugate $\overline{\omega}$.

Proposition 3.5.2. *Let G be an abelian countable group, and equip $\mathbb{C}^G$, the vector space of all functions $G \to \mathbb{C}$, with the topology of pointwise convergence, i.e. the locally convex topology induced by the family of semi-norms $(\sigma_g)_{g\in G}$, where*

$$\sigma_g(\omega) = |\omega(g)| \quad \textit{for } \omega \in \widehat{G}.$$

Then $\widehat{G}$ is a compact subspace of $\mathbb{C}^G$, and the group multiplication and group inversion of $\widehat{G}$ become continuous maps.

Proof. Since $\mathbb{T}$ is compact, the product space $\mathbb{T}^G$ is compact by Tychonoff's Theorem (Theorem 1.3.8), and $\widehat{G} \subseteq \mathbb{T}^G$. Therefore, it suffices to show that $\widehat{G}$ is closed in $\mathbb{T}^G$.

Let $(\omega_\lambda)_{\lambda\in\Lambda}$ be a net in $\widehat{G}$ that converges pointwise to $\omega \in \mathbb{T}^G$. If $g, h \in G$, then

$$\omega(gh) = \lim_\lambda \omega_\lambda(gh) = \lim_\lambda \omega_\lambda(g)\omega_\lambda(h) = (\lim_\lambda \omega_\lambda(g))(\lim_\lambda \omega_\lambda(h)) = \omega(g)\omega(h).$$

This shows that $\omega \in \widehat{G}$, so $\widehat{G}$ is a closed subspace of $\mathbb{T}^G$, hence compact.

We now show that group multiplication is continuous: Suppose $(\omega_\lambda)_\lambda$ and $(\chi_\lambda)_\lambda$ are nets in $\widehat{G}$ that converge pointwise to ω and χ, respectively. If $g \in G$ we get

$$\lim_\lambda (\omega_\lambda \chi_\lambda)(g) = \lim_\lambda \omega_\lambda(g)\chi_\lambda(g) = (\lim_\lambda \omega_\lambda(g))(\lim_\lambda \chi_\lambda(g)) = \omega(g)\chi(g) = (\omega\chi)(g).$$

This shows that the net $(\omega_\lambda\chi_\lambda)_\lambda$ converges pointwise to $\omega\chi$, which proves continuity of group multiplication. We also get for $g \in G$ that

$$\lim_\lambda \omega_\lambda^{-1}(g) = \lim_\lambda \overline{\omega_\lambda(g)} = \overline{\lim_\lambda \omega_\lambda(g)} = \overline{\omega(g)} = \omega^{-1}(g).$$

Hence group inversion is continuous as well. ■

In light of the above results, we will always view the Pontryagin dual as a compact topological space with the topology of pointwise convergence.

Example 3.5.3. (a) Let $G = \mathbb{Z}$, the additive group of integers. We can obtain characters on $\mathbb{Z}$ as follows: For any $z \in \mathbb{T}$, define $\omega_z : \mathbb{Z} \to \mathbb{T}$ by $\omega_z(k) = z^k$. Then ω_z is easily seen to be a character on $\mathbb{Z}$. Moreover, note that if $\omega \in \widehat{\mathbb{Z}}$ the $\omega(k) = \omega(1)^k$ for all $k \in \mathbb{Z}$, so the characters ω_z for $z \in \mathbb{T}$ are in fact all the characters on $\mathbb{Z}$. Moreover, the map $\widehat{\mathbb{Z}} \to \mathbb{T}$ given by $\omega \mapsto \omega(1)$ is a continuous group homomorphism, which is actually a bijection by the preceding discussion. This shows that $\widehat{\mathbb{Z}} \cong \mathbb{T}$.

(b) Let $G = \mathbb{Z}_n$, the finite cyclic group of order n. We write its elements as $[k]$ for $k \in \mathbb{Z}$, where $[k] = [l]$ if and only if $k - l \in n\mathbb{Z}$. As in (a), if $\omega \in \widehat{\mathbb{Z}_n}$ then $\omega([k]) = \omega([1])^k$ for all $[k] \in \mathbb{Z}_n$. However, since $[n] = [0]$, we need $\omega([1])^n = 1$, i.e. it has to be an nth root of unity. Thus $\widehat{\mathbb{Z}_n} \cong \{z \in \mathbb{T} : z^n = 1\}$ which is isomorphic to $\mathbb{Z}_n$.

We have that $\widehat{G} \subseteq \ell^\infty(G)$, since $|\omega(g)| = 1$ for all $\omega \in \widehat{G}$. Thus, every $\omega \in \widehat{G}$ defines a bounded linear functional on $\ell^1(G)$ via

$$\phi_\omega(a) = \sum_{g \in G} a(g)\omega(g) \quad \text{for } a \in \ell^1(G).$$

It turns out that these linear functionals belong to $\Delta(\ell^1(G))$, and that the map $\omega \mapsto \phi_\omega$ implements a homeomorphism between $\widehat{G}$ and $\Delta(\ell^1(G))$:

Theorem 3.5.4. *Let G be an abelian countable group. Then the map $f : \widehat{G} \to \Delta(\ell^1(G))$ given by*

$$f(\omega)(a) = \sum_{g \in G} a(g)\omega(g) \quad \textit{for } \omega \in \widehat{G} \textit{ and } a \in \ell^1(G)$$

is a homeomorphism.

Proof. Let $\omega \in \widehat{G}$. We must show that $f(\omega)$ is an element of $\Delta(\ell^1(G))$. We know that it is a bounded linear functional on $\ell^1(G)$ since $\omega \in \ell^\infty(G)$, but we must check that it preserves multiplication and is unital. Let $a, b \in \ell^1(G)$. Since

$$\sum_{h \in G}\sum_{g \in G} |a(h)b(h^{-1}g)\omega(g)| = \sum_{h \in G}\sum_{g \in G} |a(h)b(g)| = \Big(\sum_{h \in G} |a(h)|\Big)\Big(\sum_{g \in G} |b(g)|\Big) = \|a\|_1\|b\|_1 < \infty,$$

the interchanging of sums in the following computation is justified by Fubini's Theorem:

$$\begin{aligned}
f(\omega)(a*b) &= \sum_{g\in G}(a*b)(g)\omega(g)\\
&= \sum_{g\in G}\sum_{h\in G} a(h)b(h^{-1}g)\omega(g)\\
&= \sum_{h\in G}\sum_{g\in G} a(h)b(h^{-1}g)\omega(g)\\
&= \sum_{h\in G}\sum_{g\in G} a(h)b(g)\omega(hg)\\
&= \sum_{h\in G}\sum_{g\in G} a(h)b(g)\omega(h)\omega(g)\\
&= \Big(\sum_{h\in G} a(h)\omega(h)\Big)\Big(\sum_{g\in G} b(g)\omega(g)\Big)\\
&= f(\omega)(a)f(\omega)(b).
\end{aligned}$$

This shows that $f(\omega)$ preserves multiplication. Also,

$$f(\omega)(\delta_e) = \sum_{g\in G}\delta_e(g)\omega(g) = \omega(e) = 1,$$

which shows that $f(\omega)$ is unital. We have now shown that $f(\omega)\in\Delta(\ell^1(G))$ for every $\omega\in\widehat{G}$.

Next, we show continuity of f. Suppose $(\omega_\lambda)_{\lambda\in\Lambda}$ is a net in $\widehat{G}$ that converges to $\omega\in\widehat{G}$. We must show that the net $(f(\omega_\lambda))_\lambda$ converges to $f(\omega)$. Since the topology on $\Delta(\ell^1(G))$ is that of pointwise convergence, we must show that $(f(\omega_\lambda)(a))_\lambda$ converges to $f(\omega)(a)$ for every $a\in\ell^1(G)$, so let $a\in\ell^1(G)$ and $\epsilon>0$. Then we can find a finite subset F of G such that $\sum_{g\in G\setminus F}|a(g)|<\epsilon/4$.

Set $C=\max\{|a(g)| : g\in F\}$. Since $(\omega_\lambda)_\lambda$ converges to ω pointwise, we can find $\lambda_0\in\Lambda$ such that

$$|\omega_\lambda(g)-\omega(g)| < \frac{\epsilon}{2|F|(C+1)} \quad \text{for all } g\in F \text{ and } \lambda\geq\lambda_0.$$

Now if $\lambda\geq\lambda_0$ then

$$\begin{aligned}
|f(\omega_\lambda)(a)-f(\omega)(a)| &= \Big|\sum_{g\in G} a(g)\omega_\lambda(g) - \sum_{g\in G} a(g)\omega(g)\Big|\\
&\leq \sum_{g\in G}|a(g)||\omega_\lambda(g)-\omega(g)|\\
&= \sum_{g\in F}|a(g)||\omega_\lambda(g)-\omega(g)| + \sum_{g\in G\setminus F}|a(g)||\omega_\lambda(g)-\omega(g)|\\
&\leq \sum_{g\in F} C\cdot\frac{\epsilon}{2|F|(C+1)} + \sum_{g\in G\setminus F}|a(g)|(|\omega_\lambda(g)|+|\omega(g)|)\\
&\leq \frac{\epsilon}{2} + 2\sum_{g\in G\setminus F}|a(g)|
\end{aligned}$$

$$< \frac{\epsilon}{2} + 2 \cdot \frac{\epsilon}{4} = \epsilon.$$

This shows that f is continuous.

We show that f is injective. Let $\omega, \chi \in \widehat{G}$ and suppose that $f(\omega) = f(\chi)$. Then for every $h \in G$, we have that

$$\omega(h) = \sum_{g \in G} \delta_h(g)\omega(g) = f(\omega)(\delta_h) = f(\chi)(\delta_h) = \sum_{g \in G} \delta_h(g)\chi(g) = \chi(h).$$

Hence $\omega = \chi$.

Surjectivity of f goes as follows: Let $\tau \in \Delta(\ell^1(G))$. Then in particular $\tau \in \ell^1(G)^*$, so we know that τ is given by $\tau(a) = \sum_{g \in G} a(g)b(g)$ for some $b \in \ell^\infty(G)$. Using the multiplicativity of τ, we obtain for every $g, h \in G$ that

$$\begin{aligned} b(g)b(h) &= \Big(\sum_{k \in G} b(k)\delta_g(k)\Big)\Big(\sum_{k \in G} \delta_h(k)b(k)\Big) \\ &= \tau(\delta_g)\tau(\delta_h) \\ &= \tau(\delta_g * \delta_h) \\ &= \tau(\delta_{gh}) \\ &= \sum_{k \in G} \delta_{gh}(k)b(k) = b(gh). \end{aligned}$$

This shows that b must be multiplicative. Similarly

$$b(1) = \sum_{k \in G} \delta_e(k)b(k) = \tau(\delta_e),$$

which shows that b is unital. We know have that $b \in \widehat{G}$ and $f(b) = \tau$, which proves surjectivity.

We have now established that f is a continuous bijection from $\widehat{G}$ to $\Delta(\ell^1(G))$. Since $\widehat{G}$ is compact and $\Delta(\ell^1(G))$ is Hausdorff, it must be a homeomorphism. ■

In view of Theorem 3.5.4, let us consider the Gelfand transform of $\ell^1(G)$. It is the map $\Gamma \colon \ell^1(G) \to C(\Delta(\ell^1(G)))$ given by $\Gamma(a)(\tau) = \tau(a)$ for all $a \in \ell^1(G)$ and $\tau \in \Delta(\ell^1(G))$. Identifying $\Delta(\ell^1(G))$ with $\widehat{G}$ using the homeomorphism f from Theorem 3.5.4, the Gelfand transform is the map

$$\Gamma \colon \ell^1(G) \to C(\widehat{G})$$

given by

$$\Gamma(a)(\omega) = \sum_{g \in G} a(g)\omega(g) \quad \text{for } a \in \ell^1(G) \text{ and } \omega \in \widehat{G}.$$

This map is known as the *Fourier transform* associated with G. As an example, let us look at $G = \mathbb{Z}$. We saw that $\widehat{\mathbb{Z}} \cong \mathbb{T}$ in Example 3.5.3 (*a*), where $\omega \in \widehat{\mathbb{Z}}$ can be identified with $\omega(1) \in \mathbb{T}$. Using this identification, the Gelfand transform of $\ell^1(\mathbb{Z})$ is the map $\Gamma \colon \ell^1(\mathbb{Z}) \to C(\mathbb{T})$ given by

$$\Gamma(a)(z) = \sum_{k \in \mathbb{Z}} a(k)z^k \quad \text{for } z \in \mathbb{T}.$$

If we write $z = e^{2\pi it}$ for $t \in \mathbb{R}/\mathbb{Z}$, then we get

$$\Gamma(a)(t) = \sum_{k\in\mathbb{Z}} a(k)e^{2\pi ikt}.$$

This series is the *Fourier series* with coefficients $(a(k))_{k\in\mathbb{Z}}$.

We identify $C(\mathbb{T})$ with the continuous functions $f\colon \mathbb{R} \to \mathbb{C}$ that have period 1, i.e. $f(t+k) = f(t)$ for all $t \in \mathbb{R}$ and $k \in \mathbb{Z}$. For each $n \in \mathbb{Z}$, we let $e_n \in C(\mathbb{T})$ be the function $e_n(t) = e^{2\pi int}$. We know that $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal basis for the Hilbert space $L^2(\mathbb{T})$, with inner product given by $\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t$. The nth Fourier coefficient of a function $f \in L^2(\mathbb{T})$ is given by

$$\widehat{f}(n) = \langle f, e_n\rangle = \int_0^1 f(t)e^{-2\pi int}\,\mathrm{d}t.$$

Definition 3.5.5. The *Wiener algebra* is the set $W(\mathbb{T})$ of functions $f \in C(\mathbb{T})$ that have absolutely summable Fourier coefficients, i.e.

$$\sum_{n\in\mathbb{N}} |\widehat{f}(n)| < \infty.$$

One can show that the Wiener algebra is precisely the image $\Gamma(\ell^1(\mathbb{Z})) \subseteq C(\mathbb{T})$ under the Gelfand transform of $\ell^1(\mathbb{Z})$.

Proposition 3.5.6. *Let $f \in W(\mathbb{T})$ have the property that $f(t) \neq 0$ for all $t \in \mathbb{T}$. Then $1/f \in W(\mathbb{T})$.*

Proof. Let f be as in the proposition. Then we can write $f = \Gamma(a)$ for some $a \in \ell^1(\mathbb{Z})$. Since $f(t) \neq 0$ for all $t \in \mathbb{T}$, f is invertible as an element of $C(\mathbb{T})$, so $0 \notin \mathrm{sp}(f) = \mathrm{sp}(\Gamma(a))$. From Theorem 3.4.11 we have that $\mathrm{sp}(\Gamma(a)) = \mathrm{sp}(a)$, so a must be invertible as an element of $\ell^1(\mathbb{Z})$. But then $a * a^{-1} = \delta_0$, so

$$1 = \Gamma(\delta_0) = \Gamma(a * a^{-1}) = \Gamma(a)\Gamma(a^{-1}) = f \cdot \Gamma(a^{-1}).$$

This shows that $1/f = \Gamma(a^{-1})$, so $1/f \in W(\mathbb{T})$. ■

Remark. What we have done in this section can be generalized to *locally compact Hausdorff abelian groups* G, that is, abelian groups which carry a locally compact Hausdorff topology for which the group multiplication and group inversion are continuous maps. For these groups, one can define the Pontryagin dual $\widehat{G}$ similarly, which itself becomes a locally compact Hausdorff abelian group.

Chapter 4

C*-algebras

4.1 Involutions

Definition 4.1.1. Let A be a (complex) algebra. An *involution* on A is a map $A \to A$, $a \mapsto a^*$, satisfying the following properties:

(a) $(\lambda a + \mu b)^* = \overline{\lambda} a^* + \overline{\mu} b^*$ for all $a, b \in A$.

(b) $(ab)^* = b^* a^*$ for all $a, b \in A$.

(c) $(a^*)^* = a$ for all $a \in A$.

An algebra A equipped with an involution $*$ is called a $*$-*algebra*.

Remark. Suppose A is a unital algebra with involution $*$. Let $a \in A$. Then

$$a \cdot 1^* = (a^*)^* \cdot 1^* = (1 \cdot a^*)^* = (a^*)^* = a,$$

and similarly $1^* \cdot a = a$. By uniqueness of multiplicative identities, this implies that $1^* = 1$.

A subset S of a $*$-algebra A is called $*$-*closed* if $a^* \in S$ whenever $a \in S$. A subalgebra B of a $*$-algebra A is called a $*$-*subalgebra* if it is $*$-closed.

Definition 4.1.2. A *Banach* $*$-*algebra* is a Banach algebra A that carries an involution which is isometric with respect to the norm on A: That is, for every $a \in A$ we have that

$$\|a^*\| = \|a\|.$$

Example 4.1.3. (a) Consider $C(\Omega)$ for a compact Hausdorff space Ω. Then $f^* = \overline{f}$, i.e. pointwise complex conjugation, is an involution on $C(\Omega)$, and $C(\Omega)$ becomes a Banach $*$-algebra with respect to this involution: For $f \in C(\Omega)$ we have that

$$\|f^*\| = \sup_{t \in \Omega} |\overline{f(t)}| = \sup_{t \in \Omega} |f(t)| = \|f\|.$$

(b) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Then similarly to (a), $L^\infty(\Omega, \mathcal{A}, \mu)$ becomes a Banach $*$-algebra with respect to the involution $f^* = \overline{f}$.

(c) Consider $\mathcal{B}(H)$ for a Hilbert space H. Then the adjoint, $T \mapsto T^*$, defines an involution on $\mathcal{B}(H)$. This is known from MAT4400.

(d) Consider $\ell^1(G)$ for a group G. Given $a \in \ell^1(G)$, define

$$a^*(g) = \overline{a(g^{-1})} \quad \text{for } g \in G.$$

We leave it as an exercise to check that this defines an involution on $\ell^1(G)$, and that $\ell^1(G)$ becomes a Banach $*$-algebra with respect to this involution.

Definition 4.1.4. Let A and B be $*$-algebras. A homomorphism $\phi\colon A \to B$ is called $*$-*preserving* or a $*$-*homomorphism* if

$$\phi(a^*) = \phi(a)^* \quad \text{for all } a \in A.$$

If ϕ is additionally a bijection, then ϕ is called a $*$-*isomorphism*.

Observation 4.1.5. Let A be a unital $*$-algebra. Observe that for $a \in A$, we have that $a \in GL(A)$ if and only if $a^* \in GL(A)$, with $(a^*)^{-1} = (a^{-1})^*$: Indeed, if $a \in GL(A)$ we have that

$$a^*(a^{-1})^* = (a^{-1}a)^* = 1^* = 1,$$

and similarly $(a^{-1})^*a^* = 1$. Conversely, if $a^* \in GL(A)$, then $a = (a^*)^* \in GL(A)$ by what we already proved.

In particular, $\lambda 1 - a \in GL(A)$ if and only if $\overline{\lambda}1 - a^* = (\lambda 1 - a)^* \in GL(A)$, which shows that $\mathrm{sp}(a^*) = \overline{\mathrm{sp}(a)}$.

Definition 4.1.6. A *C*-algebra* is a Banach algebra equipped with an involution satisfying the so-called *C*-identity*:

$$\|a^*a\| = \|a\|^2 \quad \text{for all } a \in A.$$

Remark. A C*-algebra is automatically a Banach $*$-algebra: If $a \in A$, $a \neq 0$, then

$$\|a\|^2 = \|a^*a\| \leq \|a^*\|\|a\|,$$

so we can cancel $\|a\|$ to obtain $\|a\| \leq \|a^*\|$. Applying the same argument again, we get $\|a^*\| \leq \|(a^*)^*\| = \|a\|$, so $\|a\| = \|a^*\|$ for all $a \in A$ (the case $a = 0$ is trivial). Note also that in any Banach $*$-algebra we have that $\|a^*a\| \leq \|a^*\|\|a\| = \|a\|^2$, so it is the reverse inequality that distinguishes Banach $*$-algebras from C*-algebras.

Remark. If A is a unital nonzero C*-algebra, then we do not have to assume that $\|1\| = 1$: We have that

$$\|1\| = \|1^*1\| = \|1\|^2,$$

so $\|1\| = 1$ since $1 \neq 0$.

Example 4.1.7. (a) Let H be a Hilbert space. Then $\mathcal{B}(H)$ is a C*-algebra: For $x \in H$ and $T \in \mathcal{B}(H)$ we have that

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \|T^*Tx\|\|x\|.$$

Taking the supremum over all $\|x\| = 1$, we obtain $\|T\|^2 \leq \|T^*T\|$.

(b) If Ω is a compact Hausdorff space, then $C(\Omega)$ is a C*-algebra: For $f \in C(\Omega)$ we have that

$$\|f^*f\|_\infty = \|\overline{f}f\|_\infty = \||f|^2\|_\infty = \|f\|_\infty^2.$$

(c) If G is a countable group, then the Banach $*$-algebra $\ell^1(G)$ is *not* a C*-algebra in general. A concrete counter-example will be given in an exercise.

Definition 4.1.8. Let A be a unital C*-algebra. We call $a \in A$

(a) *normal* if $aa^* = a^*a$,

(b) *self-adjoint* if $a^* = a$,

(c) *unitary* if $aa^* = a^*a = 1$, i.e. $a^{-1} = a^*$.

Proposition 4.1.9. *Let A be a unital C*-algebra, and let $a \in A$. Then the following hold:*

(a) If a is unitary, then $\mathrm{sp}(a) \subseteq \mathbb{T}$.

(b) If a is self-adjoint, then $\mathrm{sp}(a) \subseteq \mathbb{R}$.

Proof. Exercise. ■

Lemma 4.1.10. *Let A be a unital C*-algebra and let a be a normal element of A. Then*

$$r(a) = \|a\|.$$

Proof. Assume first that a is self-adjoint, i.e. $a = a^*$. Applying the C*-identity, we get $\|a^2\| = \|a^*a\| = \|a\|^2$. Repeating the argument, we get $\|a^4\| = \|(a^2)^*(a^2)\| = \|a^2\|^2 = \|a\|^4$. By induction, we get $\|a^{2^n}\| = \|a\|^{2^n}$ for all $n \in \mathbb{N}$. The Spectral Radius Formula (Theorem 3.3.13) now gives

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \lim_{n\to\infty} \|a^{2^n}\|^{1/(2^n)} = \|a\|.$$

Assume now that a is normal. Then $(a^n)^*(a^n) = (a^*a)^n$, so

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \lim_{n\to\infty} \|(a^n)^*(a^n)\|^{1/(2n)} = (\lim_{n\to\infty} \|(a^*a)^n\|^{1/n})^{1/2}.$$

However, a^*a is self-adjoint, since $(a^*a)^* = a^*(a^*)^* = a^*a$. Thus, using what we already proved, the above limit is equal to

$$r(a^*a)^{1/2} = \|a^*a\|^{1/2} = \|a\|.$$

This finishes the proof. ■

Observation 4.1.11. Let A be a C*-algebra and let $a \in A$. We then define

$$\mathrm{Re}(a) = \frac{a + a^*}{2} \qquad \mathrm{Im}(a) = \frac{a - a^*}{2i}.$$

One then checks that $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ are self-adjoint, and that

$$a = \mathrm{Re}(a) + i\,\mathrm{Im}(a).$$

In fact, this representation is unique: If $b + ic = b' + ic'$ for self-adjoint elements $b, c, b', c' \in A$, then $b - b' = i(c' - c)$. By Proposition 4.1.9 and the Spectral Mapping Theorem, the spectrum of $i(c' - c)$ is a subset of $i\mathbb{R}$. On the other hand, the spectrum of $b - b'$ is a subset of $\mathbb{R}$. Thus, we must have $\operatorname{sp}(b - b') = \{0\}$. Since $b - b'$ is normal, we get $\|b - b'\| = r(b - b') = 0$, so $b = b'$. It then follows that $c = c'$.

Theorem 4.1.12. *Let A be a unital commutative C*-algebra. Then the Gelfand transform $\Gamma\colon A \to C(\Delta(A))$ is an isometric $*$-isomorphism.*

Proof. We must show that Γ is $*$-preserving, isometric and surjective.

To show that Γ is $*$-preserving, assume first that $a \in A$ is self-adjoint. Let $\tau \in \Delta(A)$. Since $\tau(a) \in \operatorname{sp}(a)$ and $\operatorname{sp}(a) \subseteq \mathbb{R}$ by Proposition 4.1.9, it follows that $\tau(a) = \overline{\tau(a)}$. Hence

$$\Gamma(a)^*(\tau) = \overline{\Gamma(a)(\tau)} = \overline{\tau(a)} = \tau(a) = \Gamma(a)(\tau) = \Gamma(a^*)(\tau).$$

Now let $a \in A$ be general. Using Observation 4.1.11 and what we already proved, we get

$$\begin{aligned}\Gamma(a)^* &= \Gamma(\operatorname{Re}(a) + i\operatorname{Im}(a))^* \\ &= \Gamma(\operatorname{Re}(a))^* - i\Gamma(\operatorname{Im}(a))^* \\ &= \Gamma(\operatorname{Re}(a)^*) - i\Gamma(\operatorname{Im}(a)^*) \\ &= \Gamma(\operatorname{Re}(a)^* - i\operatorname{Im}(a)^*) \\ &= \Gamma(a^*).\end{aligned}$$

This shows that Γ is $*$-preserving.

Next, we show that Γ is isometric. Since A is commutative, every element is normal, so Lemma 4.1.10 gives that $\|a\| = r(a)$ for all $a \in A$. Thus, using Theorem 3.4.11 (b), we get

$$\|\Gamma(a)\|_\infty = r(a) = \|a\|.$$

We now show surjectivity of Γ. Note that since Γ is an isometry, its range $B = \Gamma(A) \subseteq C(\Delta(A))$ is closed in $\Delta(A)$. Furthermore, since Γ is a unital $*$-homomorphism, B is a unital subalgebra of $C(\Delta(A))$ which is closed under conjugation. We show that B separates points: If $\tau_1, \tau_2 \in \Delta(A)$ with $\tau_1 \neq \tau_2$ then there exists $a \in A$ such that $\tau_1(a) \neq \tau_2(a)$. But then $\Gamma(a)(\tau_1) = \tau_1(a) \neq \tau_2(a) = \Gamma(a)(\tau_2)$. Since $\Gamma(a) \in B$, this shows that B separates points. By Corollary 3.2.10, it follows that $B = C(\Delta(A))$, so Γ is surjective. ■

Remark. Theorem 4.1.12 is a classification of unital, commutative C*-algebras: They are all of the form $C(\Omega)$ for some compact Hausdorff space Ω. A more general result is true, although we will not prove it here: A (possibly non-unital) commutative C*-algebra is isomorphic to $C_0(\Omega)$ for some *locally compact* Hausdorff space Ω.

4.2 The continuous functional calculus

Let A be a unital C*-algebra. A closed $*$-subalgebra of A is called a *C*-subalgebra* of A. If we have a unital C*-subalgebra B of A (i.e. the multiplicative identity of A is contained in B), then we may ask the following question: If $b \in B$ has an inverse in A, is the inverse necessarily in B? Although this property does not hold for algebra in general, it does hold for C*-algebras, as we show in the following proposition:

Proposition 4.2.1. *Let A be a unital C^*-algebra, and let B be a unital C^*-subalgebra of A. Then if $b \in B$ is invertible as an element of A, i.e. there exists $a \in A$ such that $ab = ba = 1$, then $a \in B$.*

Proof. For $x \in B$, denote by $\mathrm{sp}_B(x)$ the spectrum of x as an element of b, i.e. the set of $\lambda \in \mathbb{C}$ such that $x - \lambda 1$ does not have an inverse in B.

Assume first that b is self-adjoint. For each $n \in \mathbb{N}$, set

$$b_n = b - \frac{i}{n}1 \in B.$$

Then by the Spectral Mapping Theorem for polynomials (Proposition 3.3.11) we get

$$\mathrm{sp}_B(b_n) = \{\lambda - i/n : \lambda \in \mathrm{sp}_B(b)\}.$$

Since b is self-adjoint, $\mathrm{sp}_B(b) \subseteq \mathbb{R}$ by Proposition 4.1.9. Thus, by the above equation, $0 \notin \mathrm{sp}_B(b_n)$ for each $n \in \mathbb{N}$, so each b_n has an inverse in B. By continuity of taking inverses (Proposition 3.3.4), we have that $(b_n^{-1})_{n\in\mathbb{N}} \to b^{-1}$. Since $b_n^{-1} \in B$ for each $n \in \mathbb{N}$ and B is closed, it follows that $b^{-1} \in B$.

Next, let $b \in B$ be a general element which has an inverse in A. Then b^* is invertible as well by Observation 4.1.5, so b^*b has an inverse in A. Since b^*b is self-adjoint, the inverse of b^*b is in B by what we already proved. But then

$$b^{-1} = b^{-1}(b^*)^{-1}b^* = (b^*b)^{-1}b^* \in B.$$

This finishes the proof. ■

Remark. As a consequence, we obtain in the setting of Proposition 4.2.1

$$\mathrm{sp}_B(b) = \mathrm{sp}_A(b)$$

for every $b \in B$, where $\mathrm{sp}_B(b)$ (resp. $\mathrm{sp}_A(b)$) is the spectrum of b as an element of B (resp. of A).

Given an element $a \in A$, we wish to consider the unital C*-subalgebra $C^*(a)$ generated by a, i.e. the least unital C*-subalgebra of A that contains a. We can describe this C*-subalgebra as the intersection of all unital C*-subalgebras of A that contain a (here we use the fact that intersections of $*$-algebras are $*$-algebras, and that intersections of closed sets are closed).

When a is normal, we can obtain a simple description of a. Note that a product of any finite number of copies of a's and a^*'s should be in $C^*(a)$. Since a and a^* commute, we can collect powers in such a product, and represent it as $a^m(a^*)^n$ for positive integers m and n. Complex linear combinations of such products are exactly of the form $p(a, a^*)$, where $p(x, y) \in \mathbb{C}[x, y]$ is a complex polynomial in two variables x and y. One verifies that $\{p(x, y) : x, y \in \mathbb{C}[x, y]\}$ is a $*$-subalgebra of A, and that its closure is equal to $C^*(a)$. From this it also follows that $C^*(a)$ is commutative.

Theorem 4.2.2 (Continuous Functional Calculus)**.** *Let A be a unital C^*-algebra, and let $a \in A$ be normal. Then there exists a unique bounded unital $*$-homomorphism*

$$C(\mathrm{sp}(a)) \to A,$$

$$f \mapsto f(a),$$

such that $\iota(a) = a$*, where* ι *is the inclusion of* $\mathrm{sp}(a)$ *into* $\mathbb{C}$*. Moreover, the homomorphism has the following properties:*

(a) It is isometric, i.e. $\|f(a)\| = \|f\|_\infty$ *for every* $f \in C(\mathrm{sp}(a))$*,*

(b) the image of the map $C(\mathrm{sp}(a)) \to A$ *is exactly the* C^**-subalgebra of* A *generated by* a*.*

Proof. We begin by showing existence. Let $B = C^*(a)$, the unital C^*-subalgebra generated by a. Since a is normal, B is commutative, so we consider the character space $\Delta(B)$. By Proposition 3.4.8 (b), $\mathrm{sp}_B(a) = \{\tau(a) : \tau \in \Delta(B)\}$. By Proposition 4.2.1, $\mathrm{sp}_B(a) = \mathrm{sp}_A(a)$ (we denote the latter simply by $\mathrm{sp}(a)$). Thus, we can define a map $h\colon \Delta(B) \to \mathrm{sp}(a)$ by $h(\tau) = \tau(a)$ which is surjective. We will show that h is a homeomorphism.

Since the topology on $\Delta(B)$ is that of pointwise convergence, h is continuous. Furthermore, suppose $h(\tau_1) = h(\tau_2)$ for $\tau_1, \tau_2 \in \Delta(B)$, then $\tau_1(a) = \tau_2(a)$. By Theorem 4.1.12, every character is a $*$-homomorphism, so $\tau_1(a^*) = \overline{\tau_1(a)} = \overline{\tau_2(a)} = \tau_2(a^*)$ as well. Since τ_1 and τ_2 agree on the generating set $\{a, a^*\}$ of A, it follows that $\tau_1 = \tau_2$. This shows that h is injective. We have now shown that h is a continuous bijection between compact Hausdorff spaces, so it follows from the usual topological argument that h must be a homeomorphism.

We consider now the Gelfand transform $\Gamma\colon B \to C(\Delta(B))$ of B. By Theorem 4.1.12, Γ is an isometric $*$-isomorphism, so its inverse is also an isometric $*$-isomorphism. Denote by $\phi\colon C(\mathrm{sp}(a)) \to C(\Delta(B))$ the map given by $\phi(f) = f \circ h$. By an exercise, ϕ is a unital isometric $*$-isomorphism since h is a homeomorphism. We thus have a map $\psi\colon C(\mathrm{sp}(a)) \to B$ given by the composition $\Gamma^{-1} \circ \phi$ which is a unital isometric $*$-isomorphism. Furthermore, the function $\iota \circ h\colon \Delta(B) \to \mathbb{C}$ is given by $(\iota \circ h)(\tau) = \iota(\tau(a)) = \tau(a)$ for $\tau \in \Delta(B)$. Hence $\iota \circ h = \Gamma(a)$, so $\psi(\iota) = \Gamma^{-1}(\iota \circ h) = a$. This proves the existence of a unital $*$-homomorphism $C(\mathrm{sp}(a)) \to A$ such that $\iota \mapsto a$, and also shows that it is necessarily isometric and that its image is B.

Finally, we show uniqueness. Suppose $\psi_1, \psi_2\colon C(\mathrm{sp}(a)) \to A$ are two $*$-homomorphisms such that $\psi_1(\iota) = \psi_2(\iota)$. Since ι separates the points of $\mathrm{sp}(a)$, the C^*-subalgebra generated by ι is all of $C(\mathrm{sp}(a))$. Therefore, since ψ_1, ψ_2 are $*$-homomorphisms that agree on ι, they must agree on all of $C(\mathrm{sp}(a))$. ■

Example 4.2.3 (Diagonal operators)**.** Let H be a separable, infinite-dimensional Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{N}}$. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a bounded sequence of complex numbers. Then we can define an operator $D \in \mathcal{B}(H)$ by

$$De_n = \lambda_n e_n \quad \text{for } n \in \mathbb{N}.$$

Then each λ_n is an eigenvalue for D, so $\mathrm{Cl}\{\lambda_n : n \in \mathbb{N}\} \subseteq \mathrm{sp}(D)$. On the other hand, if $\lambda \notin \mathrm{Cl}\{\lambda_n : n \in \mathbb{N}\}$, then $|\lambda - \lambda_n| \geq c > 0$ for all $n \in \mathbb{N}$. Hence the sequence $(1/(\lambda_n - \lambda))_{n\in\mathbb{N}}$ is bounded, so $Re_n = (\lambda_n - \lambda)^{-1} e_n$ defines a bounded linear operator which is the inverse of $D - \lambda I$. This shows that $\mathrm{sp}(D) = \mathrm{Cl}\{\lambda_n : n \in \mathbb{N}\}$.

One checks that D^* is determined by $D^* e_n = \overline{\lambda_n} e_n$. The operator D^*D is then determined by $D^*De_n = |\lambda_n|^2 e_n$, which shows that D is normal. Given $f \in C(\mathrm{sp}(D))$, we can now ask: What is $f(D)$? We claim that $f(D)$ is determined by

$$f(D)e_n = f(\lambda_n)e_n \quad \text{for } n \in \mathbb{N}.$$

To show this, we observe first that since f is continuous and $(\lambda_n)_n$ is bounded, the sequence $(f(\lambda_n))_n$ is also bounded, so we can define an operator D_f by $D_f e_n = f(\lambda_n)e_n$. One now checks that the map $C(\mathrm{sp}(D)) \to \mathcal{B}(H)$, $f \mapsto D_f$ is a unital $*$-homomorphism. Furthermore, $D_\iota e_n = \iota(\lambda_n)e_n = \lambda_n e_n = De_n$, so we have $D_f = f(D)$ by uniqueness of the continuous functional calculus.

Proposition 4.2.4 (Spectral Mapping Theorem)**.** *Let a be a normal element in a unital C^*-algebra A, and let $f \in C(\mathrm{sp}(a))$. Then*

$$\mathrm{sp}(f(a)) = f(\mathrm{sp}(a)).$$

Proof. We use the fact that $C(\mathrm{sp}(a)) \to B = C^*(a)$, $f \mapsto f(a)$ is an isometric $*$-isomorphism:

$$\mathrm{sp}(f(a)) = \mathrm{sp}_B(f(a)) = \mathrm{sp}_{C(\mathrm{sp}(a))}(f) = f(\mathrm{sp}(a)).$$

■

Proposition 4.2.5. *Let a be a normal element in a unital C^*-algebra A. If $f \in C(\mathrm{sp}(a))$ and $g \in C(\mathrm{sp}(f(a)))$, then*

$$(g \circ f)(a) = g(f(a)).$$

Proof. Exercise. ■

Proposition 4.2.6. *Let A be a unital C^*-algebra, and let $a \in A$ be normal. Then the following hold:*

(a) a is unitary if and only if $\mathrm{sp}(a) \subseteq \mathbb{T}$.

(b) a is self-adjoint if and only if $\mathrm{sp}(a) \subseteq \mathbb{R}$.

Proof. We have already shown the forward impliations in Proposition 4.1.9. Let $a \in A$ be normal, and let $\iota \in C(\mathrm{sp}(a))$ be the inclusion $\mathrm{sp}(a) \to \mathbb{C}$. Then ι^* is given by $\iota^*(z) = \overline{z}$ for $z \in \mathrm{sp}(a)$. Since $C(\mathrm{sp}(a)) \to A$, $f \mapsto f(a)$ is a $*$-homomorphism with $\iota(a) = a$, we have that $\iota^*(a) = \iota(a)^* = a^*$.

(a): If $\mathrm{sp}(a) \subseteq \mathbb{T}$, then

$$(\iota \cdot \iota^*)(z) = \iota(z)\iota^*(z) = z\overline{z} = 1$$

for all $z \in \mathrm{sp}(a)$, so $\iota \cdot \iota^* = 1$. It follows that

$$aa^* = \iota(a)\iota^*(a) = 1(a) = 1.$$

Since a is normal, we also have that $a^*a = 1$, so a is unitary.

(b): If $\mathrm{sp}(a) \subseteq \mathbb{R}$, then

$$\iota^*(z) = \overline{z} = z = \iota(z) \quad \text{for all } z \in \mathrm{sp}(a),$$

so $\iota^* = \iota$. It then follows that $a^* = \iota^*(a) = \iota(a) = a$. ■

Example 4.2.7. Let a be a normal element of a unital C*-algebra, and suppose $\mathrm{sp}(a)$ is disconnected: Specifically, let C be a subset of $\mathrm{sp}(a)$ which is both open and closed, and not equal to $\emptyset$ or $\mathrm{sp}(a)$. Then the indicator function $\mathbb{1}_C$ is continuous on $\mathrm{sp}(a)$, so we can define $p = \mathbb{1}_C(a)$. Since $\mathbb{1}_C$ is real-valued, p is self-adjoint. Furthermore, $\mathbb{1}_C^2 = \mathbb{1}_C$, so $p^2 = p$. We call elements satisfying $p^* = p = p^2$ *orthogonal projections.* In any C*-algebra, 0 and 1 are orthogonal projections. Since C is neither $\emptyset$ nor $\mathrm{sp}(a)$, $\mathbb{1}_C$ is neither 0 nor 1, so p is neither 0 nor 1.

As a particular example, consider the C*-algebra $\mathrm{Mat}_n(\mathbb{C})$ of complex $n \times n$ matrices. Let $M \in \mathrm{Mat}_n(\mathbb{C})$ be a normal matrix. Let $\mathrm{sp}(M) = \{\lambda_1, \ldots, \lambda_k\}$, where $k \leq n$. Here, each $\{\lambda_i\}$ is clopen, so the functions $f_i = \mathbb{1}_{\{\lambda_i\}}$ are continuous on $\mathrm{sp}(M)$. Note also that $\sum_{i=1}^k f_i = 1$, $f_i f_j = 0$ if $i \neq j$ and $\sum_{i=1}^k \lambda_i f_i = \iota$. Now set

$$P_i = f_i(M).$$

By the continuous functional calculus, we have that $P_i P_j = 0$ for $i \neq j$, $\sum_{i=1}^n P_i = I$ and $\sum_{i=1}^n \lambda_i P_i = M$.

4.3 Positivity

Definition 4.3.1. Let A be a unital C*-algebra. A self-adjoint element $a \in A$ is called *positive* if $\mathrm{sp}(a) \subseteq [0, \infty)$. We write $a \geq 0$ to indicate this. We write A_+ for the subset of positive elements of A.

Example 4.3.2. (a) Consider the C*-algebra $\mathbb{C}$. For $z \in \mathbb{C}$ we have that $\mathrm{sp}(z) = \{z\}$, so z is positive if and only if z is a nonnegative number.

(b) Consider the C*-algebra $C(\Omega)$ for a compact Hausdorff space Ω, and let $f \in C(\Omega)$. Since $\mathrm{sp}(f) = \mathrm{Im}(f)$, we have that f is positive precisely when $\mathrm{Im}(f) \subseteq [0, \infty)$, i.e. when $f(t) \geq 0$ for all $t \in \Omega$.

Proposition 4.3.3. *Let A be a unital C*-algebra and let $a \in A$. Then the following are equivalent:*

(a) a is positive.

(b) $a = b^2$ for some self-adjoint $b \in A$.

Moreover, if a is positive, then there exists a unique positive $b \in A$ such that $b^2 = a$, and we write $b = a^{1/2}$.

Proof. (a) $\Rightarrow$ (b): Suppose a is positive. Since $\mathrm{sp}(a) \subseteq [0, \infty)$, we can define a function $f\colon \mathrm{sp}(a) \to \mathbb{C}$ by $f(z) = z^{1/2}$. Then $f^2(z) = f(z)^2 = z = \iota(z)$, i.e. $f^2 = \iota$. We also have that $f^* = f$. Set $b = f(a)$ according to the continuous functional calculus. Then b is self-adjoint, since $b^* = f(a)^* = f^*(a) = f(a) = b$. Furthermore, by the spectral mapping theorem (Proposition 4.2.4) we get

$$\mathrm{sp}(b) = f(\mathrm{sp}(a)) = \{\lambda^{1/2} : \lambda \in \mathrm{sp}(a)\} \subseteq [0, \infty),$$

so b is in fact positive. Finally

$$b^2 = f(a)^2 = f^2(a) = \iota(a) = a.$$

$(b) \Rightarrow (a)$: Suppose $a = b^2$ for some self-adjoint $b \in A$. Then $\mathrm{sp}(a) = \mathrm{sp}(b^2) = \{\lambda^2 : \lambda \in \mathrm{sp}(b)\}$. Since b is self-adjoint, $\mathrm{sp}(b) \subseteq \mathbb{R}$, so $\mathrm{sp}(a) \subseteq [0, \infty)$.

As we saw above, if a is positive, we can find a positive $b \in A$ such that $b^2 = a$, namely $b = f(a)$. Suppose now that c is positive and $c^2 = a$. Since c commutes with a and $b \in C^*(a)$, c must commute with b. Let B be the unital C*-algebra generated by b and c, which is commutative. Let $\Gamma\colon B \to C(\Delta(B))$ be the Gelfand transform of B. Then $\Gamma(b)^2 = \Gamma(b^2) = \Gamma(c^2) = \Gamma(c)^2$. Since b and c are positive, $\Gamma(b)$ and $\Gamma(c)$ are functions on $\Delta(B)$ that take nonnegative values. It follows that $\Gamma(b) = \Gamma(c)$, so $b = c$ by injectivity of the Gelfand transform. ■

Proposition 4.3.4. *Suppose a and b are positive elements of a unital C*-algebra A. Then $a + b$ is positive.*

Proof. Let a and b be positive elements of A. If $\lambda \in \mathrm{sp}(a)$, then $\lambda \geq 0$, so $\lambda \leq \|a\| \leq \|a\| + \lambda$. Hence $0 \leq \|a\| - \lambda \leq \|a\|$, so by normality of $a - \|a\|1$ and the spectral mapping theorem for polynomials we get

$$\|a - \|a\|1\| = r(a - \|a\|1) = \sup_{\lambda \in \mathrm{sp}(a)} |\lambda - \|a\|| = \sup_{\lambda \in \mathrm{sp}(a)} (\|a\| - \lambda) \leq \|a\|.$$

Similarly $\|b - \|b\|1\| \leq \|b\|$. Now if $\lambda \in \mathrm{sp}(a+b)$, then $\lambda - \|a\| - \|b\| \in \mathrm{sp}(a + b - (\|a\| + \|b\|)1)$, so

$$\begin{aligned} |\lambda - \|a\| - \|b\|| &\leq \|(a+b) - (\|a\| + \|b\|)1\| \\ &\leq \|a - \|a\|1\| + \|b - \|b\|1\| \\ &\leq \|a\| + \|b\|. \end{aligned}$$

But then $\lambda - \|a\| - \|b\| \geq -\|a\| - \|b\|$, so $\lambda \geq 0$. This proves that $\mathrm{sp}(a + b) \subseteq [0, \infty)$, so $a + b$ is positive. ■

We define two functions $f_+, f_- \colon \mathbb{R} \to \mathbb{C}$ by

$$f_+(t) = \begin{cases} t & \text{if } t \geq 0, \\ 0 & \text{if } t < 0 \end{cases} \qquad f_-(t) = \begin{cases} 0 & \text{if } t \geq 0, \\ -t & \text{if } t < 0 \end{cases}.$$

Let $\iota\colon \mathbb{R} \to \mathbb{C}$ denote the function $\iota(t) = t$ for $t \in \mathbb{R}$. Note that $\iota = f_+ - f_-$, $f_+ f_- = 0$ and that $\mathrm{abs} = f_+ + f_-$, where $\mathrm{abs}(t) = |t|$.

Definition 4.3.5. Let A be a unital C*-algebra, and let $a \in A$ be self-adjoint. We then define

$$\begin{aligned} a_+ &= f_+(a), \\ a_- &= f_-(a), \\ |a| &= \mathrm{abs}(a). \end{aligned}$$

Remark. Note that since a is self-adjoint, $\mathrm{sp}(a) \subseteq \mathbb{R}$, so the definitions in Definition 4.3.5 make sense since f_+, f_- and abs are defined on $\mathbb{R}$. By the continuous functional calculus, a_+, a_- and $|a|$ are positive elements of A since f_+, f_- and the absolute value function are nonnegative.

From the continuous functional calculus, we get the following identities for a self-adjoint:

$$\begin{aligned} a &= a_+ - a_-, \\ |a| &= a_+ + a_- = (a^2)^{1/2}, \\ a_+a_- &= 0. \end{aligned}$$

Lemma 4.3.6. *Let A be a unital algebra and let $a, b \in A$. Then $1 - ab$ is invertible if and only $1 - ba$ is invertible, and consequently*

$$\mathrm{sp}(ab) \setminus \{0\} = \mathrm{sp}(ba) \setminus \{0\}.$$

Proof. Exercise. ■

Proposition 4.3.7. *An element a of a unital C*-algebra is positive if and only if $a = b^*b$ for some $b \in A$.*

Proof. If a is positive, then by Proposition 4.3.3 we can find a self-adjoint b such that $b^2 = a$. Hence $b^*b = b^2 = a$.

To prove the converse, we show first that if $b \in A$ and $-b^*b$ is positive, then $b = 0$. By Lemma 4.3.6, if $-b^*b$ is positive, then $-bb^*$ is positive. Write $b = c + id$ where c and d are self-adjoint as in Observation 4.1.11. Then $b^*b + bb^* = 2c^2 + 2d^2$, so

$$b^*b = (2c^2 + 2d^2) + (-bb^*)$$

which is positive by Proposition 4.3.3 and Proposition 4.3.4. Hence both b^*b and $-b^*b$ are positive, so $\mathrm{sp}(b^*b) \subseteq [0, \infty) \cap (-\infty, 0] = \{0\}$. Since b^*b is normal, this implies $\|b\|^2 = \|b^*b\| = r(b^*b) = 0$, i.e. $b = 0$.

Let now $b \in A$ be general and set $a = b^*b$. Then a is self-adjoint, so we can write $a = a_+ - a_-$ where a_+ and a_- are positive and $a_+a_- = a_-a_+ = 0$ as in Definition 4.3.5. Set $c = ba_-$. Then

$$-c^*c = -a_-^*b^*ba_- = -a_-^*(a_+ - a_-)a_- = a_-^3 \in A_+.$$

Hence, by what we already proved, we must have $c = 0$. But then $ba_- = 0$, so

$$0 = b^*ba_- = aa_- = (a_+ - a_-)a_- = -a_-^2.$$

Thus, $\|a_-\|^2 = \|a_-^*a_-\| = \|a_-^2\| = 0$, so $a_- = 0$. This shows that $a = a_+ \in A_+$, so a is positive. ■

Definition 4.3.8. Let a be an element of a unital C*-algebra A. We then define the *absolute value* of a to be

$$|a| = (a^*a)^{1/2}.$$

If $a, b \in A$ are self-adjoint, we write $a \leq b$ if $b - a \geq 0$, that is, $b - a$ is a positive element.

Proposition 4.3.9. *The following hold in a unital C*-algebra A:*

(a) If $a, b, c \in A$ are self-adjoint and $a \leq b$, then $a + c \leq b + c$.

(b) If $a, b \in A$ be self-adjoint, $a \leq b$ and $c \in A$, then $c^*ac \leq c^*bc$.

(c) If a and b are positive and $a \leq b$, then $\|a\| \leq \|b\|$.

(d) If a is positive, then $a \leq \|a\|1$.

(e) If a and b are positive and invertible, and $a \leq b$, then $b^{-1} \leq a^{-1}$.

Proof. Exercise. ■

4.4 Partial isometries and polar decomposition

Proposition 4.4.1. *Let u be an element of a unital C*-algebra A. Then the following are equivalent:*

(a) $u = uu^*u$.

(b) $u^* = u^*uu^*$.

(c) u^*u is an orthogonal projection.

(d) uu^* is an orthogonal projection.

Proof. $(a) \Leftrightarrow (b)$: If $u = uu^*u$, then by conjugating both sides we get $u^* = u^*uu^*$. Conversely, if we start with $u^* = u^*uu^*$, then we conjugate again and get back to $u = uu^*u$.

$(b) \Rightarrow (c)$: If $u^* = u^*uu^*$, then

$$(u^*u)^2 = u^*uu^*u = u^*u.$$

Since u^*u is self-adjoint, this shows that u^*u is an orthogonal projection.

$(c) \Rightarrow (a)$: Suppose that $p = u^*u$ is an orthogonal projection. Set $a = uu^*u - u$. Then

$$a^*a = p^3 - p^2 - p^2 + p = 0.$$

Hence $\|a^*a\| = 0$, so $\|a\|^2 = \|a^*a\| = 0$. This shows that $uu^*u = u$.

$(b) \Leftrightarrow (d)$: This follows from the equivalence of (a) and (c) by interchanging u with u^*. ■

Definition 4.4.2. We call an element in a unital C*-algebra A satisfying any of the equivalent conditions in Proposition 4.4.1 a *partial isometry*.

Note that u is a partial isometry if and only if u^* is a partial isometry from the first two conditions of Proposition 4.4.1.

We will now characterize partial isometries in the C*-algebra $\mathcal{B}(H)$, where H is a Hilbert space. This characterization explains the name partial isometry.

Proposition 4.4.3. *Let H be a Hilbert space and let $U \in \mathcal{B}(H)$. Then the following are equivalent:*

(a) U is a partial isometry in $\mathcal{B}(H)$.

(b) There exists a closed subspace M of H such that $U|_M$ is an isometry and $U|_{M^\perp} = 0$.

Furthermore, if U is a partial isometry, then U has closed range, the above subspace M is uniquely determined and equals $(\operatorname{Ker} U)^\perp$, and U^*U is the orthogonal projection onto M, while UU^* is the orthogonal projection onto $N = U(M)$.

Proof. Suppose that U is a partial isometry. By Proposition 4.4.1 (c), $P = U^*U$ is an orthogonal projection. Set $M = \operatorname{Im} P$, which is a closed subspace of H. Let $x \in M$, so that $Px = x$. Then by Proposition 4.4.1 (d), we get

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle U^*Ux, x\rangle = \langle x, x\rangle = \|x\|^2.$$

This shows that $U|_M$ is an isometry. On the other hand, if $x \in M^\perp = \operatorname{Ker} P$, then

$$\|Ux\|^2 = \langle U^*Ux, x\rangle = 0,$$

so $U|_{M^\perp} = 0$.

Conversely, suppose that there exists a closed subspace M of H such that $U|_M$ is an isometry and $U|_{M^\perp} = 0$. Let P be the orthogonal projection onto M. If $x \in M$, then

$$\langle U^*Ux, x\rangle = \|Ux\|^2 = \|x\|^2 = \langle Px, x\rangle.$$

On the other hand, if $x \in M^\perp$, then

$$\langle U^*Ux, x\rangle = \langle U^*0, x\rangle = 0 = \langle Px, x\rangle.$$

Thus, $\langle U^*Ux, x\rangle = \langle Px, x\rangle$ for all $x \in H$, so $U^*U = P$. By Proposition 4.4.1 (c), this shows that U is a partial isometry, and that U^*U is the orthogonal projection onto M. Since $\operatorname{Ker} U = \operatorname{Ker} U^*U = \operatorname{Ker} P = M^\perp$, we get $M = (\operatorname{Ker} U)^\perp$, which shows that M is uniquely determined.

We leave the proof that U has closed range as an exercise. Because of this, $N = U(H) = U(M)$ is closed. Set $Q = UU^*$. If $y \in H$, then $Q(Uy) = (UU^*)Uy = (UU^*U)y = Uy$, so $N \subseteq \operatorname{Im} Q$. Conversely, if $x \in \operatorname{Im} Q$, then $UU^*x = x$. If $y \in \operatorname{Ker} U^*$, then

$$\langle x, y\rangle = \langle UU^*x, y\rangle = \langle U^*x, U^*y\rangle = 0.$$

Consequently $x \in (\operatorname{Ker} U^*)^\perp = U(H) = N$. This shows that UU^* is the orthogonal projection onto N. ■

Let U be a partial isometry. We call $M = U^*U(H)$ the *initial space* of U, and $N = UU^*(H)$ the *final space* of U. From Proposition 4.4.3, we have that $U|_M$ is an isometry, while $U|_{M^\perp} = 0$. Thus, the restricted and corestricted map $U|_M \colon M \to N$ is unitary. Conversely, if $V \colon M \to N$ is a unitary map, we can extend it to a partial isometry $\tilde{V} \colon H \to H$ with initial space M and final space N by setting

$$\tilde{V}(x) = \begin{cases} Vx & \text{if } x \in M, \\ 0 & \text{if } x \in M^\perp. \end{cases}$$

If U is a partial isometry on H with initial space M and final space N, then U^* is a partial isometry with initial space N and final space M: Indeed, the projection onto the initial space of U^* is $(U^*)^*U^* = UU^*$, which is exactly the projection onto the final space of U, and similarly for the final space of U^*.

Theorem 4.4.4 (Polar decomposition)**.** *Let H be a Hilbert space, and let $T \in \mathcal{B}(H)$.*

(a) There is a unique positive operator $P \in \mathcal{B}(H)$ such that

$$\|Tx\| = \|Px\| \quad \text{for all } x \in H,$$

and this operator is $P = |T|$.

(b) There is a unique partial isometry $U \in \mathcal{B}(H)$ such that

$$T = U|T| \quad \text{and} \quad \operatorname{Ker} U = \operatorname{Ker} T,$$

The initial space of U is $(\operatorname{Ker} T)^\perp$ and the final space of U is $\operatorname{Cl}(\operatorname{Im} T)$.

Proof. (a): First, we note that

$$\||T|x\|^2 = \langle |T|^*|T|x, x\rangle = \langle T^*Tx, x\rangle = \|Tx\|^2$$

for any $x \in H$. This shows that $|T|$ has the required property. We now show uniqueness: Suppose $S \in \mathcal{B}(H)$, $S \geq 0$, and $\|Sx\| = \|Tx\|$ for all $x \in H$. Then

$$\langle S^2x, x\rangle = \|Sx\|^2 = \|Tx\|^2 = \langle T^*Tx, x\rangle$$

for all $x \in H$. Thus $S^2 = T^*T$. Since S^2 is positive, S is its unique square root, so $S = |T|$.

(b): To begin with, suppose $x, x' \in H$ are such that $|T|x = |T|x'$. Then $|T|(x - x') = 0$, so by what (a) we get $\|Tx - Tx'\| = \||T|(x - x')\| = 0$, i.e., $Tx = Tx'$. This shows that the map $U_0 \colon \operatorname{Im}|T| \to \operatorname{Im} T$ given by

$$U_0(|T|x) = Tx \quad \text{for } x \in H$$

is well-defined. We leave it as an exercise to show that U_0 is linear. Furthermore,

$$\|U_0(|T|x)\| = \|Tx\| = \||T|x\| \quad \text{for } x \in H,$$

which shows that U_0 is an isometry. Thus, U_0 extends uniquely to a linear isometry $\operatorname{Cl}(\operatorname{Im}|T|) \to \operatorname{Cl}(\operatorname{Im} T)$, which we also denote by U_0. Set $M = \operatorname{Cl}(\operatorname{Im}|T|)$. Then $M^\perp = (\operatorname{Im}|T|)^\perp = \operatorname{Ker}|T|^* = \operatorname{Ker}|T| = \operatorname{Ker} T$, where the last equality follows from (a). Hence we can define a partial isometry $U \colon H \to H$ by

$$Ux = \begin{cases} U_0x & \text{if } x \in \operatorname{Cl}(\operatorname{Im}|T|), \\ 0 & \text{if } x \in \operatorname{Ker} T. \end{cases}$$

The initial space of U is then $M = (\operatorname{Ker} T)^\perp$. Moreover, we get

$$U|T|x = U_0|T|x = Tx \quad \text{for all } x \in H,$$

so $U|T| = T$. Furthermore, $\operatorname{Ker} U = M^\perp = \operatorname{Ker} T$. The final space of U is $U(M) = \operatorname{Cl}\{U|T|x : x \in H\} = \operatorname{Cl}\{Tx : x \in H\} = \operatorname{Cl}(\operatorname{Im} T)$.

Finally, we show uniqueness. Suppose $V \in \mathcal{B}(H)$ is a partial isometry that satisfies $T = V|T|$ and $\operatorname{Ker} V = \operatorname{Ker} T$. Then $V|T|x = U|T|x$ for all $x \in H$, so $V|_M = U|_M$. Moreover, $V|_{M^\perp} = V|_{\operatorname{Ker} T} = V|_{\operatorname{Ker} V} = 0$. Hence $V = U$, which proves the uniqueness part. ∎

We call the decomposition $T = U|T|$ of $T \in \mathcal{B}(H)$ the *polar decomposition* of T.

Example 4.4.5. Let $H = \mathbb{C}$, so that $\mathcal{B}(H) \cong \mathbb{C}$. If $z \in \mathbb{C}$ then $|z| = (z^*z)^{1/2}$ is the usual absolute value of z. A partial isometry is a complex number w such that $w^*w = |w|^2$ is a projection, i.e., $w = 0$ or $|w| = 1$. In the latter case we can write $w = e^{i\theta}$ for some $\theta \in \mathbb{R}$. Thus, the polar decomposition of a nonzero complex number is $z = e^{i\theta}|z|$ for θ an argument of z.

Example 4.4.6. Let H be a finite-dimensional Hilbert space, and let $T \in \mathcal{B}(H)$. Let $T = U|T|$ be the polar decomposition of T. Then U is a unitary map from $(\operatorname{Ker} T)^\perp$ to $\operatorname{Im} T$. Since $H = \operatorname{Im} T + (\operatorname{Im} T)^\perp$, we have that $\dim H = \dim \operatorname{Im} T + \dim(\operatorname{Im} T)^\perp$. By the rank-nullity theorem, we also have $\dim H = \dim \operatorname{Im} T + \dim \operatorname{Ker} T$, so $\dim \operatorname{Ker} T = \dim(\operatorname{Im} T)^\perp$. Thus, we can find a unitary map $V\colon \operatorname{Ker} T \to (\operatorname{Im} T)^\perp$. Define $\tilde{U} \in \mathcal{B}(H)$ by

$$\tilde{U}(x) = \begin{cases} Ux & \text{if } x \in (\operatorname{Ker} T)^\perp, \\ Vx & \text{if } x \in \operatorname{Ker} T. \end{cases}$$

Since U is a unitary from $(\operatorname{Ker} T)^\perp$ to $\operatorname{Im} T$ and V is a unitary from $\operatorname{Ker} T$ to $(\operatorname{Im} T)^\perp$, it follows that $\tilde{U}$ is a unitary from H to H. Moreover, $\tilde{U}|T|x = U|T|x = Tx$ for all $x \in H$, so $\tilde{U}|T| = T$. This is often referred to as a polar decomposition of T, but note that the unitary map $\tilde{U}$ is not unique unless T is invertible.

Proposition 4.4.7. *Let $T \in \mathcal{B}(H)$, and let $T = U|T|$ be the polar decomposition of T. Then the following identities hold:*

$$\begin{aligned} U^*U|T| &= |T| \\ UU^*T &= T \\ |T^*| &= U|T|U^* \\ T^* &= U^*|T^*|. \end{aligned}$$

Proof. In general, we have the identity $\operatorname{Cl}(\operatorname{Im} T) = \operatorname{Cl}(\operatorname{Im} TT^*)$ for $T \in \mathcal{B}(H)$. For the operator $|T|$, this gives

$$\operatorname{Cl}(\operatorname{Im} |T|) = \operatorname{Cl}(\operatorname{Im} |T||T|^*) = \operatorname{Cl}(\operatorname{Im} |T|^2) = \operatorname{Cl}(\operatorname{Im} T^*T) = \operatorname{Cl}(\operatorname{Im} T^*).$$

Hence $(\operatorname{Ker} T)^\perp = \operatorname{Cl}(\operatorname{Im} T^*) = \operatorname{Cl}(\operatorname{Im} |T|)$, so the initial space of U equals $\operatorname{Cl}(\operatorname{Im} |T|)$. Since U^*U is the orthogonal projection onto the initial space, it follows that $U^*U|T| = |T|$.

The rest of the identities are left as an exercise. ∎

4.5 Normal operators

Recall that a bounded linear operator T on a Hilbert space H is called *normal* if it is normal as an element of the C*-algebra $\mathcal{B}(H)$, i.e., $TT^* = T^*T$. We begin with the following characterization of normal operators.

Proposition 4.5.1. *Let $T \in \mathcal{B}(H)$. Then the following are equivalent:*

(a) T is normal.

(b) $\|Tx\| = \|T^*x\|$ for all $x \in H$.

Proof. For $x \in H$ we have that

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle,$$

while

$$\|T^*x\|^2 = \langle T^*x, T^*x \rangle = \langle TT^*x, x \rangle.$$

Thus, condition (b) is equivalent to

$$\langle T^*Tx, x \rangle = \langle TT^*x, x \rangle \quad \text{for all } x \in H.$$

This is again equivalent to $T^*T = TT^*$, i.e., normality of T. ∎

Let H be a Hilbert space. Recall that a bounded linear operator T on H is called *bounded below* if there exists a constant $\alpha > 0$ such that

$$\alpha\|x\| \leq \|Tx\| \quad \text{for all } x \in H.$$

Recall that T is bounded below if and only if T is injective and has closed range. Consequently, T is invertible if and only if T is bounded below and has dense range. If both T and T^* are bounded below, then $\operatorname{Im} T$ is closed, $\operatorname{Ker}(T) = \{0\}$ and $\operatorname{Im}(T)^\perp = \operatorname{Ker}(T^*) = \{0\}$; hence $\operatorname{Im}(T) = \operatorname{Cl}(\operatorname{Im}(T)) = H$, so T is invertible.

Definition 4.5.2. Let T be a bounded, linear operator on a Hilbert space H. A number $\lambda \in \mathbb{C}$ is called an *approximate eigenvalue* of T if there exists a sequence $(x_n)_{n\in\mathbb{N}}$ of unit vectors in H such that

$$\lim_{n\to\infty} (T - \lambda I)x_n = 0.$$

Remark. (a) Every eigenvalue λ of T is an approximate eigenvalue of T: Simply take a corresponding eigenvector x of norm 1 and consider the constant sequence $(x)_{n\in\mathbb{N}}$. Then $\lim_{n\to\infty}(T - \lambda I)x = 0$.

(b) Every approximate eigenvalue of T belongs to the spectrum of T, as can be seen by a contrapositive argument: If $T - \lambda I$ is invertible and $(x_n)_{n\in\mathbb{N}}$ is any sequence of unit vectors in H, then

$$1 = \|x_n\| = \|(T - \lambda I)^{-1}(T - \lambda I)x_n\| \leq \|(T - \lambda I)^{-1}\|\|(T - \lambda I)x_n\|,$$

which shows that $(T - \lambda I)x_n$ cannot converge to 0 as $n \to \infty$.

Proposition 4.5.3. *Let H be a Hilbert space, let $T \in \mathcal{B}(H)$ and let $\lambda \in \mathbb{C}$. Then the following are equivalent:*

(a) $\lambda \in \operatorname{sp}(T)$.

(b) *λ is an approximate eigenvalue of T, or $\overline{\lambda}$ is an approximate eigenvalue of T^*.*

Proof. (*a*) ⇒ (*b*): Suppose $\lambda \in \operatorname{sp}(T)$, so that $S = \lambda I - T$ is not invertible. Then either S or S^* is not bounded below: If S is not bounded below, then for every $n \in \mathbb{N}$ we can find $y_n \in H$ such that

$$\frac{1}{n}\|y_n\| > \|Sy_n\|.$$

Letting $x_n = \|y_n\|^{-1}y_n$, we get $\|Sx_n\| < 1/n$ for each $n \in \mathbb{N}$, which shows that $S(x_n)_n \to 0$. Hence λ is an approximate eigenvalue for T. Similarly, if $S^* = \overline{\lambda}I - T^*$ is not bounded below, then $\overline{\lambda}$ is an approximate eigenvalue of T^*.

(*b*) ⇒ (*a*): We have seen that an approximate eigenvalue of T is in the spectrum of T. If $\overline{\lambda}$ is an approximate eigenvalue of T^*, then $\overline{\lambda} \in \operatorname{sp}(T^*) = \overline{\operatorname{sp}(T)}$, so $\lambda \in \operatorname{sp}(T)$. ∎

Corollary 4.5.4. *Suppose $T \in \mathcal{B}(H)$ is normal. Then every $\lambda \in \operatorname{sp}(T)$ is an approximate eigenvalue of T.*

Proof. If $T \in \mathcal{B}(H)$ is normal and $\lambda \in \mathbb{C}$, then $T - \lambda I$ is normal, so by Proposition 4.5.1 we get

$$\|(T - \lambda I)x\| = \|(T^* - \overline{\lambda}I)x\| \quad \text{for all } x \in H. \tag{4.1}$$

If $\lambda \in \operatorname{sp}(T)$ then by Proposition 4.5.3 either λ is an approximate eigenvalue of T or $\overline{\lambda}$ is an approximate eigenvalue of T^*. If $\overline{\lambda}$ is an approximate eigenvalue of T^*, then we can find a sequence $(x_n)_{n\in\mathbb{N}}$ of unit vectors in H such that $(T^* - \overline{\lambda}I)x_n \to 0$ as $n \to \infty$. By (4.1), this implies that $(T - \lambda I)x_n \to 0$ as $n \to \infty$, so in either case, λ is an approximate eigenvalue of T. ∎

Proposition 4.5.5. *Let H be a Hilbert space, and let $T \in \mathcal{B}(H)$. Then the following are equivalent:*

(a) T is positive (as an element of the C-algebra $\mathcal{B}(H)$).*

(b) $\langle Tx, x\rangle \geq 0$ for all $x \in H$.

Proof. (*a*) ⇒ (*b*). Suppose that T is positive. By Proposition 4.3.7, we can write $T = S^*S$ for some $S \in \mathcal{B}(H)$. We then get for $x \in H$ that

$$\langle Tx, x\rangle = \langle S^*Sx, x\rangle = \langle Sx, Sx\rangle = \|Sx\|^2 \geq 0.$$

(*b*) ⇒ (*a*): Suppose that $\langle Tx, x\rangle \geq 0$ for all $x \in H$. Then in particular, $\langle Tx, x\rangle$ is real, so we get

$$\langle T^*x, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle} = \langle Tx, x\rangle.$$

Since this holds for all $x \in H$, we have that $T^* = T$, so T is in particular normal. If $\lambda \in \operatorname{sp}(T)$, then λ is an approximate eigenvalue of T by Corollary 4.5.4, so we can find a sequence $(x_n)_{n\in\mathbb{N}}$ of unit vectors in H such that $(T - \lambda I)x_n \to 0$ as $n \to \infty$. Then

$$|\lambda - \langle Tx_n, x_n\rangle| = |\langle(\lambda I - T)x_n, x_n\rangle| \leq \|(\lambda I - T)x_n\|\|x_n\| = \|(\lambda I - T)x_n\| \to 0.$$

Thus $\lambda = \lim_{n\to\infty}\langle Tx_n, x_n\rangle \geq 0$. Since $\lambda \in \operatorname{sp}(T)$ was arbitrary, we conclude that T is positive. ∎

From the above proposition it follows that $S \leq T$ for self-adjoint operators $S, T \in \mathcal{B}(H)$ if and only if $\langle Sx, x\rangle \leq \langle Tx, x\rangle$ for all $x \in H$.

Let T be a bounded, linear operator on a Hilbert space H. We say that $\lambda \in \operatorname{sp}(T)$ is *isolated* if there exists $\delta > 0$ such that $\operatorname{sp}(T) \cap B_\delta(\lambda) = \{\lambda\}$, i.e., $\{\lambda\}$ is a clopen subset of $\operatorname{sp}(T)$. As we saw in Example 4.2.7, the characteristic function $\mathbb{1}_{\{\lambda\}}$ is then continuous on $\operatorname{sp}(T)$, so we can form the orthogonal projection $P = \mathbb{1}_{\{\lambda\}}(T)$.

Proposition 4.5.6. *Let $T \in \mathcal{B}(H)$ be a normal operator, and suppose $\lambda \in \operatorname{sp}(T)$ is isolated. Then λ is an eigenvalue of T, and $P = \mathbb{1}_{\{\lambda\}}(T)$ is the orthogonal projection onto the eigenspace* $\operatorname{Ker}(\lambda I - T)$ *of λ.*

Proof. As usual, denote by ι the inclusion of $\operatorname{sp}(T)$ into $\mathbb{C}$. Consider the function $(\lambda\cdot 1-\iota)\cdot\mathbb{1}_{\{\lambda\}}$ on $\operatorname{sp}(T)$. This function is identically zero on $\operatorname{sp}(T)$, which implies that $(\lambda I - T)P = 0$. Thus, if $Px = x$, then

$$(\lambda I - T)x = (\lambda I - T)(Px) = 0x = 0,$$

which shows that $\operatorname{Im} P \subseteq \operatorname{Ker}(\lambda I - T)$. To show the reverse inclusion, we define a function $g\colon \operatorname{sp}(T) \to \mathbb{C}$ by

$$g(z) = \begin{cases} (\lambda - z)^{-1} & \text{if } z \neq \lambda, \\ 0 & \text{if } z = \lambda. \end{cases}$$

Since λ is isolated in $\operatorname{sp}(T)$, g is continuous. Moreover,

$$g \cdot (\lambda \cdot 1 - \iota) = \mathbb{1}_{\operatorname{sp}(T)\setminus\{\lambda\}} = 1 - \mathbb{1}_{\{\lambda\}},$$

so $g(T)(\lambda I - T) = I - P$. Hence, if $x \in \operatorname{Ker}(\lambda I - T)$, then $(\lambda I - T)x = 0$, so

$$0 = g(T)(\lambda I - T)x = (I - P)x = x - Px.$$

Hence $Px = x$, so $x \in \operatorname{Im} P$. This shows that $\operatorname{Ker}(\lambda I - T) \subseteq \operatorname{Im} P$, so P is the orthogonal projection onto $\operatorname{Ker}(\lambda I - T)$. Since $\mathbb{1}_{\{\lambda\}} \neq 0$, it follows that $P \neq 0$. Hence $\operatorname{Ker}(\lambda I - T)$ is not the zero subspace, so λ is an eigenvalue of T. ∎

4.6 The Borel functional calculus

We need some results concerning sesquilinear forms.

Definition 4.6.1. Let X be a complex vector space. A *sesquilinear form* on X is a map $L\colon X \times X \to \mathbb{C}$ which is linear in the first argument and conjugate-linear in the second, that is,

$$L(\lambda x + \mu y, z) = \lambda L(x, z) + \mu L(y, z),$$
$$L(x, \lambda y + \mu z) = \overline{\lambda} L(x, y) + \overline{\mu} L(x, z).$$

for all $x, y, z \in X$ and $\lambda, \mu \in \mathbb{C}$.

Example 4.6.2. (a) If X is a complex vector space, then an inner product on V is an example of a sesquilinear form.

(b) If L is a sesquilinear form on X, then $L^*: X \times X \to \mathbb{C}$ given by

$$L^*(x,y) = \overline{L(y,x)} \quad \text{for } x, y \in X,$$

is a sesquilinear form on L called the *adjoint of* L. We say that L is *self-adjoint* if $L^* = L$.

Observation 4.6.3. If L is a sesquilinear form on X, then the polarization identity holds for L just as for inner products:

$$L(x,y) = \frac{1}{4}\sum_{k=0}^{3} i^k L(x+i^k y, x+i^k y) \quad \text{for all } x, y \in X.$$

Consequently, if L and L' are sesquilinear forms on X such that $L(x,x) = L'(x,x)$ for all $x \in X$, then $L = L'$.

Definition 4.6.4. We say that a sesquilinear form L on a vector space X is *bounded* if there exists $C \geq 0$ such that

$$|L(x,y)| \leq C\|x\|\|y\| \quad \text{for all } x, y \in X.$$

If L is bounded, we define its norm by

$$\|L\| = \sup\{|L(x,y)| : x, y \in H, \|x\| \leq 1, \|y\| \leq 1\}.$$

Example 4.6.5. Let H be a Hilbert space and let $T \in \mathcal{B}(H)$. Then we can define a bounded, sesquilinear form L_T on H by

$$L_T(x,y) = \langle Tx, y\rangle \quad \text{for all } x, y \in H\ .$$

The boundedness of L follows from the Cauchy–Schwarz inequality:

$$|L_T(x,y)| = |\langle Tx, y\rangle| \leq \|Tx\|\|y\| \leq \|T\|\|x\|\|y\|.$$

This shows that $\|L_T\| \leq \|T\|$. On the other hand,

$$\|Tx\| \leq \sup_{\|y\|=1} |\langle Tx, y\rangle| \leq \sup_{\|y\|=1, \|x\|=1} |L_T(x,y)| = \|L_T\|.$$

This shows that $\|L_T\| = \|T\|$. Note also that $L_T^* = L_{T^*}$. Thus, L_T is self-adjoint if and only if T is self-adjoint.

Proposition 4.6.6. *Let H be a Hilbert space, and let L be a bounded, sesquilinear form on H. Then there exists a unique $T \in \mathcal{B}(H)$ such that $L = L_T$.*

Proof. Let $y \in H$. Define $\phi_y : H \to \mathbb{C}$ by

$$\phi_y(x) = L(x,y) \quad \text{for } x \in H.$$

Then ϕ_y is a bounded linear functional on H, since

$$|\phi_y(x)| = |L(x,y)| \leq \|L\|\|x\|\|y\|.$$

By Riesz' representation theorem, there exists a unique vector $Sy \in H$ such that

$$L(x,y) = \phi_y(x) = \langle x, Sy\rangle \quad \text{for all } x \in H,$$

with $\|\phi_y\| = \|Sy\|$. We claim that the mapping $H \to H$, $y \mapsto Sy$, defines a bounded linear map on H. Indeed, for $x, y, y' \in H$ and $\lambda, \mu \in \mathbb{C}$ we get

$$\begin{aligned} \langle x, S(\lambda y + \mu y')\rangle &= L(x, \lambda y + \mu y') \\ &= \lambda L(x,y) + \overline{\mu} L(x,y') \\ &= \lambda \langle x, Sy\rangle + \overline{\mu}\langle x, Sy'\rangle \\ &= \langle x, \lambda Sy + \mu Sy'\rangle. \end{aligned}$$

Hence $S(\lambda y + \mu y') = \lambda S(y) + \mu S(y')$. Boundedness follows from

$$\|Sy\| = \|\phi_y\| = \sup_{\|x\|\le 1} |L(x,y)| \le \|L\|\|y\|.$$

Letting T be the adjoint of S, we have that

$$L(x,y) = \langle Tx, y\rangle = L_T(x,y) \quad \text{for all } x, y \in H.$$

■

Given $T \in \mathcal{B}(H)$, we denote by $M(\mathrm{sp}(T))$ the complex vector space of all complex regular Borel measures on $\mathrm{sp}(T)$.

Proposition 4.6.7. *Let $T \in \mathcal{B}(H)$ be normal and let $x, y \in H$. Then there exists a unique $\mu_{x,y} \in M(\mathrm{sp}(T))$ called the* spectral measure of T associated with x, y *such that*

$$\langle f(T)x, y\rangle = \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y} \quad \textit{for all } f \in C(\mathrm{sp}(T)).$$

We also have that the total variation norm of $\mu_{x,y}$ satisfies $\|\mu_{x,y}\| \le \|x\|\|y\|$.

Proof. Let $x, y \in H$, and define a function $\phi_{x,y} \colon C(\mathrm{sp}(T)) \to \mathbb{C}$ by

$$\phi_{x,y}(f) = \langle f(T)x, y\rangle \quad \text{for } x, y \in H.$$

Since $f \mapsto f(T)$ is linear, it follows that $\phi_{x,y}$ is linear. We also have that

$$|\phi_{x,y}(f)| \le \|f(T)\|\|x\|\|y\| = \|f\|_\infty \|x\|\|y\|,$$

which shows that $\phi_{x,y}$ is a bounded linear functional on $C(\mathrm{sp}(T))$, with $\|\phi_{x,y}\| \le \|x\|\|y\|$. By the Riesz–Markov–Kakutani representation theorem for $C(\mathrm{sp}(T))^*$, it follows that there exists a unique $\mu_{x,y} \in M(\mathrm{sp}(T))$ such that

$$\langle f(T)x, y\rangle = \phi_{x,y}(f) = \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y} \quad \text{for all } f \in C(\mathrm{sp}(T)).$$

From the same theorem we also have that

$$\|\mu_{x,y}\| = \|\phi_{x,y}\| \le \|x\|\|y\|.$$

■

Proposition 4.6.8. *Let $T \in \mathcal{B}(H)$ be normal. For $x, y, z \in H$ and $\lambda, \gamma \in \mathbb{C}$ we have the following relations of spectral measures of T:*

(a) $\mu_{\lambda x+\gamma y,z} = \lambda\mu_{x,z} + \gamma\mu_{y,z}$.

(b) $\mu^*_{x,y} = \mu_{y,x}$, *where* $\mu^*(B) = \overline{\mu(B)}$ *for Borel sets* $B \subseteq \mathrm{sp}(T)$.

(c) $\mu_{x,x}$ *is a nonnegative measure.*

(d) $\mu_{x,y}(\mathrm{sp}(T)) = \langle x, y\rangle$. *Consequently, if* $\mu_{x,x} = 0$, *then* $x = 0$.

Proof. For all $f \in C(\mathrm{sp}(T))$ we have that

$$\begin{aligned}\int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{\lambda x+\gamma y,z} &= \langle f(T)(\lambda x + \gamma y), z\rangle \\ &= \lambda\langle f(T)x, y\rangle + \gamma\langle f(T)x, z\rangle \\ &= \lambda \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y} + \gamma \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,z} \\ &= \int_{\mathrm{sp}(T)} f \,\mathrm{d}(\lambda\mu_{x,y} + \gamma\mu_{x,z}).\end{aligned}$$

This proves (a). For part (b), we have for $f \in C(\mathrm{sp}(T))$ that

$$\begin{aligned}\int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu^*_{x,y} = \overline{\int_{\mathrm{sp}(T)} \overline{f} \,\mathrm{d}\mu_{x,y}} &= \overline{\langle \overline{f}(T)x, y\rangle} \\ = \overline{\langle f(T)^*x, y\rangle} &= \overline{\langle x, f(T)y\rangle} = \langle f(T)y, x\rangle.\end{aligned}$$

To show that $\mu_{x,x}$ is nonnegative, it suffices by the Riesz–Markov–Kakutani theorem to show that $\phi_{x,x}$ is a positive linear functional. If $f \in C(\mathrm{sp}(T))$ is such that $f(t) \geq 0$ for all $t \in \mathrm{sp}(T)$, then $f(T)$ is a positive operator since $\mathrm{sp}(f(T)) = f(\mathrm{sp}(T))$. Hence

$$\phi_{x,x}(f) = \langle f(T)x, x\rangle \geq 0,$$

so $\mu_{x,x}$ is a nonnegative measure. Finally,

$$\mu_{x,y}(\mathrm{sp}(T)) = \int_{\mathrm{sp}(T)} 1 \,\mathrm{d}\mu_{x,y} = \langle 1(T)x, y\rangle = \langle Ix, y\rangle = \langle x, y\rangle.$$

Thus, if $\mu_{x,x} = 0$ then $\|x\|^2 = \mu_{x,x}(\mathrm{sp}(T)) = 0$ so $x = 0$. ■

Let $T \in \mathcal{B}(H)$ be a normal operator. Let $B_b(\mathrm{sp}(T))$ denote the vector space of complex-valued, bounded, Borel measurable functions on $\mathrm{sp}(T)$. This becomes a unital, commutative C*-algebra with respect to the usual pointwise operations and the supremum norm. The C*-algebra $C(\mathrm{sp}(T))$ is a unital C*-subalgebra of $B_b(\mathrm{sp}(T))$.

Theorem 4.6.9 (The Borel Functional Calculus)**.** *Let $T \in \mathcal{B}(H)$ be normal. Then there exists a unique unital $*$-homomorphism $B_b(\mathrm{sp}(T)) \to \mathcal{B}(H)$, $f \mapsto f(T)$, which coincides with the continuous functional calculus when restricted to $C(\mathrm{sp}(T))$ and has the following continuity* *property: Whenever* $(f_n)_{n\in\mathbb{N}}$ *is a sequence in* $B_b(\mathrm{sp}(T))$ *with* $\sup_{n\in\mathbb{N}}\|f_n\|_\infty < \infty$ *that converges pointwise to a function* $f \in B_b(\mathrm{sp}(T))$*, then*

$$\lim_{n\to\infty} f_n(T)x = f(T)x \quad \text{for all } x \in H.$$

Moreover, the $*$*-homomorphism has the following properties:*

(a) *It is bounded, with norm equal to one, i.e.,* $\|f(T)\| \le \|f\|_\infty$ *for all* $f \in B_b(\mathrm{sp}(T))$.

(b) *If* $x, y \in H$ *and* $f \in B_b(\mathrm{sp}(T))$*, then*

$$\langle f(T)x, y\rangle = \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y}.$$

(c) *If* $S \in \mathcal{B}(H)$ *and* $ST = TS$*, then* $Sf(T) = f(T)S$ *for all* $f \in B_b(\mathrm{sp}(T))$.

Proof. Given $f \in B_b(\mathrm{sp}(T))$, we define a function $L\colon H \times H \to \mathbb{C}$ by

$$L(x, y) = \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y} \quad \text{for } x, y \in H.$$

We will show that L is bounded, sesquilinear form. Sesquilinearity follows at once from Proposition 4.6.8 (a) and (b). Also, by Proposition 4.6.8 (d), we get

$$|L(x, y)| \le \int_{\mathrm{sp}(T)} |f| \,\mathrm{d}|\mu_{x,y}| \le |\mu_{x,y}|(\mathrm{sp}(T))\|f\|_\infty = \|\mu_{x,y}\|\|f\|_\infty \le \|x\|\|y\|\|f\|_\infty.$$

Hence $\|L\| \le \|f\|_\infty$, so L is bounded. It follows from Proposition 4.6.6 that there exists a unique $f(T) \in \mathcal{B}(H)$ such that

$$\langle f(T)x, y\rangle = L(x, y) = \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y} \quad \text{for all } x, y \in H.$$

We also get $\|f(T)\| = \|L\| \le \|f\|_\infty$. Note that by Proposition 4.6.7, the map $B_b(\mathrm{sp}(T)) \to \mathcal{B}(H)$, $f \mapsto f(T)$, extends the continuous functional calculus.

We show that the map $B_b(\mathrm{sp}(T)) \to \mathcal{B}(H)$, $f \mapsto f(T)$, is a $*$-homomorphism. For linearity, let $f, g \in B_b(\mathrm{sp}(T))$ and $\lambda, \gamma \in \mathbb{C}$. Then

$$\begin{aligned}
\langle(\lambda f + \gamma g)(T)x, y\rangle &= \int_{\mathrm{sp}(T)} (\lambda f + \gamma g) \,\mathrm{d}\mu_{x,y} \\
&= \lambda \int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{x,y} + \gamma \int_{\mathrm{sp}(T)} g \,\mathrm{d}\mu_{x,y} \\
&= \lambda\langle f(T)x, y\rangle + \gamma\langle g(T)x, y\rangle \\
&= \langle(\lambda f(T) + \gamma g(T))x, y\rangle.
\end{aligned}$$

Since this holds for all $x, y \in H$, we conclude that $(\lambda f + \gamma g)(T) = \lambda f(T) + \gamma g(T)$.

Next, we show that $f \mapsto f(T)$ is $*$-preserving: If $x, y \in H$ and $f \in B_b(\mathrm{sp}(T))$, then

$$\langle \overline{f}(T)x, y\rangle = \int_{\mathrm{sp}(T)} \overline{f} \,\mathrm{d}\mu_{x,y}$$

$$
\begin{aligned}
&= \overline{\int_{\mathrm{sp}(T)} f \,\mathrm{d}\mu_{y,x}} \\
&= \overline{\langle f(T)y, x\rangle} \\
&= \langle x, f(T)y\rangle = \langle f(T)^* x, y\rangle.
\end{aligned}
$$

Since $x, y \in H$ were arbitrary, we conclude that $\overline{f}(T) = f(T)^*$.

We now show multiplicativity, which is more involved. Let $x, y \in H$ and $f, g \in B_b(\mathrm{sp}(T))$. If f and g are both continuous, then we know that $(fg)(T) = f(T)g(T)$ since the continuous functional calculus is multiplicative. In terms of spectral measures, we get

$$
\begin{aligned}
\int_{\mathrm{sp}(T)} g \,\mathrm{d}(f\mu_{x,y}) &= \int_{\mathrm{sp}(T)} fg \,\mathrm{d}\mu_{x,y} = \langle (fg)(T)x, y\rangle \\
&= \langle f(T)g(T)x, y\rangle = \langle g(T)x, f(T)^* y\rangle \\
&= \int_{\mathrm{sp}(T)} g \,\mathrm{d}\mu_{x,f(T)^*y}.
\end{aligned}
$$

Hence $\mu_{x,f(T)^*y} = f\mu_{x,y}$. But then for any $g \in B_b(\mathrm{sp}(T))$ we get

$$
\langle (fg)(T)x, y\rangle = \int_{\mathrm{sp}(T)} fg \,\mathrm{d}\mu_{x,y} = \int_{\mathrm{sp}(T)} g \,\mathrm{d}\mu_{x,f(T)^*y} = \langle f(T)g(T)x, y\rangle.
$$

Hence $(fg)(T) = f(T)g(T)$ for all $f \in C(\mathrm{sp}(T))$ and $g \in B_b(\mathrm{sp}(T))$. If instead $f \in B_b(\mathrm{sp}(T))$ and $g \in C(\mathrm{sp}(T))$, we get

$$
(fg)(T) = (fg)(T)^{**} = (\overline{fg}(T))^* = (\overline{g}\overline{f})(T)^* = (\overline{g}(T)\overline{f}(T))^* = f(T)g(T).
$$

As we have already seen, this can be expressed in terms of spectral measures as $f\mu_{x,y} = \mu_{x,f(T)^*y}$. But this time $f \in B_b(\mathrm{sp}(T))$, so if $g \in B_b(\mathrm{sp}(T))$ we get

$$
\langle (fg)(T)x, y\rangle = \int_{\mathrm{sp}(T)} g \,\mathrm{d}f\mu_{x,y} = \int_{\mathrm{sp}(T)} g \,\mathrm{d}\mu_{x,f(T)^*y} = \langle f(T)g(T)x, y\rangle.
$$

Hence $(fg)(T) = f(T)g(T)$ for all $f, g \in B_b(\mathrm{sp}(T))$.

To show the continuity property, let $(f_n)_{n\in\mathbb{N}}$ be a sequence converging to f as in the theorem. We have that

$$
|f_n - f|^2 \leq (|f_n| + |f|)^2 \leq (\sup_{n\in\mathbb{N}} \|f_n\|_\infty + \|f\|_\infty)^2 < \infty.
$$

This shows that the sequence of functions $(f_n - f)_{n\in\mathbb{N}}$ is uniformly bounded. Since each measure $\mu_{x,x}$ is finite, we can use the dominated convergence theorem as follows:

$$
\lim_{n\to\infty} \|f_n(T)x - f(T)x\|^2 = \lim_{n\to\infty} \langle |f_n(T) - f(T)|^2 x, x\rangle = \lim_{n\to\infty} \int_{\mathrm{sp}(T)} |f_n - f|^2 \,\mathrm{d}\mu_{x,x,} = 0.
$$

Hence $\lim_{n\to\infty} f_n(T)x = f(T)x$ for all $x \in H$.

The uniqueness and (c) will be given as an exercise. ■

4.7 Projection-valued measures

Proposition 4.7.1. *Let $T \in \mathcal{B}(H)$ be normal, and denote by $\mathcal{B}$ the σ-algebra of Borel sets of* $\mathrm{sp}(T)$. *Define a function $P\colon \mathcal{B} \to \mathcal{B}(H)$ given by*

$$P(B) = \mathbb{1}_B(T) \quad \text{for } B \in \mathcal{B}.$$

Then P satisfies the following properties:

(a) *$P(B)$ is an orthogonal projection for all $B \in \mathcal{B}$.*

(b) *$P(\emptyset) = 0$ and $P(\mathrm{sp}(T)) = 1$.*

(c) *$P(B \cap B') = P(B)P(B')$ for all Borel sets $B, B' \in \mathcal{B}$.*

(d) *If $(B_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint sets in $\mathcal{B}$, then*

$$P\Big(\bigcup_{n\in\mathbb{N}} B_n\Big)x = \sum_{n\in\mathbb{N}} P(B_n)x \quad \text{for all } x \in H.$$

(e) *$\langle P(B)x, y\rangle = \mu_{x,y}(B)$ for all $B \in \mathcal{B}$ and $x, y \in H$, where $\mu_{x,y}$ is the spectral measure of T associated with x and y.*

Proof. (a): This follows from applying the Borel functional calculus and the identities $\mathbb{1}_B^* = \mathbb{1}_B$ and $\mathbb{1}_B^2 = \mathbb{1}_B$.

(b): Since $\mathbb{1}_\emptyset = 0$ and $\mathbb{1}_{\mathrm{sp}(T)} = 1$, we get $P(\emptyset) = 0(T) = 0$ and $P(\mathrm{sp}(T)) = 1(T) = 1$.

(c): Since $\mathbb{1}_{B\cap B'} = \mathbb{1}_B \mathbb{1}_{B'}$, we get

$$P(B \cap B') = \mathbb{1}_{B\cap B'}(T) = (\mathbb{1}_B \mathbb{1}_{B'})(T) = \mathbb{1}_B(T)\mathbb{1}_{B'}(T) = P(B)P(B').$$

(d): Set $A = \bigcup_{n\in\mathbb{N}} A_n$ and $f_n = \sum_{k=1}^n \mathbb{1}_{B_k}$ for each $n \in \mathbb{N}$. Then $(f_n)_{n\in\mathbb{N}}$ converges pointwise to $\mathbb{1}_A$. Since the sets $(B_n)_n$ are pairwise disjoint, $|f_n| \leq 1$ for all $n \in \mathbb{N}$. We can thus use the continuity property of Theorem 4.6.9 to conclude that

$$\lim_n \sum_{k=1}^n \mathbb{1}_{B_k}(T) = \lim_n f_n(T)x = f(T)x = \mathbb{1}_B(T)x \quad \text{for all } x \in H.$$

(e): If $x, y \in H$, then

$$\langle P(B)x, y\rangle = \langle \mathbb{1}_B(T)x, y\rangle = \int_{\mathrm{sp}(T)} \mathbb{1}_B \,\mathrm{d}\mu_{x,y} = \mu_{x,y}(B).$$

■

Definition 4.7.2. The map $P\colon \mathcal{B} \to \mathcal{B}(H)$ in Proposition 4.7.1 is called the *projection-valued measure of T.*

Given $T \in \mathcal{B}(H)$ normal with associated projection-valued measure P, a common notation is to write

$$f(T) = \int_{\mathrm{sp}(T)} f \,\mathrm{d}P = \int_{\mathrm{sp}(T)} f(z)\,\mathrm{d}P(z) \quad \text{for } f \in B_b(\mathrm{sp}(T)).$$

With this notation, we get the formulas

$$T = \int_{\mathrm{sp}(T)} \iota \,\mathrm{d}P = \int_{\mathrm{sp}(T)} z\,\mathrm{d}P(z)$$
$$I = \int_{\mathrm{sp}(T)} 1\,\mathrm{d}P.$$

Compare with the formulas for a normal matrix in Example 4.2.7.

We can use P to describe the spectrum and the eigenvalues of T:

Proposition 4.7.3. *Let $T \in \mathcal{B}(H)$ be normal, and let P be the projection-valued measure of T. Given $\lambda \in \mathbb{C}$, the following hold:*

(a) $\lambda \in \mathrm{sp}(T)$ if and only if $P(B_\epsilon(\lambda) \cap \mathrm{sp}(T)) \neq 0$ for all $\epsilon > 0$.

(b) λ is an eigenvalue of T if and only if $P(\{\lambda\}) \neq 0$.

Proof. We prove (b) and leave (a) as an exercise. Let λ be an eigenvalue of T. For each $n \in \mathbb{N}$, set $E_n = \{z \in \mathrm{sp}(T) : |z-\lambda| > 1/n\}$. Then

$$\bigcup_{n\in\mathbb{N}} E_n = \mathrm{sp}(T) \setminus \{\lambda\}.$$

Let $f_n(z) = (z-\lambda)^{-1}\mathbb{1}_{E_n}(z)$. Then $f_n \in B_b(\mathrm{sp}(T))$ for each $n \in \mathbb{N}$. Since $f_n \cdot (\iota - \lambda 1) = \mathbb{1}_{E_n}$, it follows that

$$P(E_n) = \mathbb{1}_{E_n}(T) = f_n(T)(T-\lambda I).$$

Hence, if $x \in \mathrm{Ker}(T-\lambda I)$, then $P(E_n)x = 0$. Since $E_n \subseteq E_{n+1}$ for each n and $\bigcup_n E_n = \mathrm{sp}(T) \setminus \{\lambda\}$, continuity of measure gives

$$x - P(\{\lambda\})x = P(\mathrm{sp}(T)\setminus\{\lambda\})x = \lim_{n\to\infty} P(E_n)x = 0$$

for x an eigenvector of T corresponding to λ. Thus $x = P(\{\lambda\})x$, so $P(\{\lambda\}) \neq 0$ since $x \neq 0$.

Conversely, suppose $P(\{\lambda\}) \neq 0$. Then there exists $x \neq 0$ such that $P(\{\lambda\})x = x$, so

$$Tx = TP(\{\lambda\})x = (\iota\mathbb{1}_{\{\lambda\}})(T)x = \lambda\mathbb{1}_{\{\lambda\}}(T)x = \lambda P(\{\lambda\})x = \lambda x.$$

Hence x is an eigenvalue of T. In fact, we have shown that $P(\{\lambda\})$ is the projection onto the eigenspace of λ. ■

Proposition 4.7.4. *Let H be a Hilbert space. Then the linear span of the set of all the orthogonal projections in $\mathcal{B}(H)$ is dense in $\mathcal{B}(H)$.*

Proof. Let $T \in \mathcal{B}(H)$. We want to approximate T with a linear combination of projections. Since we can write $T = \mathrm{Re}(T) + i\,\mathrm{Im}(T)$ where $\mathrm{Re}(T)$ and $\mathrm{Im}(T)$ are self-adjoint (in particular normal), it suffices to find an approximation in the case where T is normal. So we assume that T is normal and let $\epsilon > 0$. Let ι denote the inclusion $\mathrm{sp}(T) \to \mathbb{C}$. Since $\iota \in B_b(\mathrm{sp}(T))$, we can find a simple function $f \in B_b(\mathrm{sp}(T))$ such that $\|\iota - f\|_\infty < \epsilon$. We can assume that f is of the form

$$f = \sum_{j=1}^{n} \lambda_j \mathbb{1}_{B_j}$$

where $\lambda_j \in \mathbb{C}$ and $B_j \subseteq \mathrm{sp}(T)$ is Borel for each j, with $B_j \cap B_k = \emptyset$ for $j \neq k$ and $\bigcup_{j=1}^{n} B_j = \mathrm{sp}(T)$. Set $P_j = P_j(B_j)$. Then

$$\left\| T - \sum_{j=1}^{n} \lambda_j P_j \right\| = \|(\iota - f)(T)\| \leq \|\iota - f\|_\infty < \epsilon.$$

This finishes the proof. ■

Chapter 5

Operators on Hilbert spaces

5.1 The trace

For a complex $n \times n$ matrix

$$M = \begin{pmatrix} m_{1,1} & \cdots & m_{1,n} \\ \vdots & \ddots & \vdots \\ m_{n,1} & \cdots & m_{n,n} \end{pmatrix},$$

the trace is defined as the sum of the diagonal entries: $m_{1,1} + \cdots m_{n,n}$. Alternatively, we can describe the trace as follows: Let $\{e_1, \ldots, e_n\}$ be the standard basis. Then for $1 \leq i, j \leq n$, we have that $m_{i,j} = \langle Me_i, e_j \rangle$. Hence

$$\operatorname{tr}(M) = \sum_{j=1}^{n} \langle Me_j, e_j \rangle.$$

This motivates the following definition:

Definition 5.1.1. Let H be a Hilbert space, and choose an orthonormal basis $(e_j)_{j \in J}$ for H. We define the *trace* of a positive operator $T \in \mathcal{B}(H)$ to be the number

$$\operatorname{tr}(T) = \sum_{j \in J} \langle Te_j, e_j \rangle.$$

Remark. Note that because T is positive, the numbers $\langle Te_j, e_j \rangle$ are nonnegative, which makes the sum appearing in the definition of $\operatorname{tr}(T)$ a well-defined, possibly infinite number. We will eventually extend it to a suitable class of operators known as the trace class operators. We will also very soon see that the definition does not depend on the choice of orthonormal basis.

Proposition 5.1.2. *If $T \in \mathcal{B}(H)$, then*

$$\operatorname{tr}(T^*T) = \operatorname{tr}(TT^*).$$

Proof. First, using Parseval's identity we get for all $i \in J$ that

$$\langle T^*Te_i, e_i \rangle = \|Te_i\|^2 = \sum_{j \in J} |\langle Te_i, e_j \rangle|^2.$$

On the other hand, for $j \in J$ we get

$$\langle TT^*e_j, e_j\rangle = \|T^*e_j\|^2 = \sum_{i\in J} |\langle T^*e_j, e_i\rangle|^2.$$

Now $|\langle Te_i, e_j\rangle|^2 = |\langle e_i, T^*e_j\rangle|^2 = |\langle T^*e_j, e_i\rangle|^2$ for all $i, j \in J$. This gives us

$$\begin{aligned}\operatorname{tr}(TT^*) &= \sum_{j\in J}\langle TT^*e_j, e_j\rangle = \sum_{j\in J}\sum_{i\in J} |\langle T^*e_j, e_i\rangle|^2 \\ &= \sum_{i\in J}\sum_{j\in J} |\langle Te_i, e_j\rangle|^2 = \sum_{i\in J}\langle T^*Te_i, e_i\rangle = \operatorname{tr}(T^*T).\end{aligned}$$

■

Corollary 5.1.3. *Let $T \in \mathcal{B}(H)_+$ and let $U \in \mathcal{B}(H)$ be unitary. Then*

$$\operatorname{tr}(UTU^*) = \operatorname{tr}(T).$$

In particular, the definition of the trace does not depend on the chosen orthonormal basis.

Proof. Using Proposition 5.1.2 and the existence of the square root of the positive operator T, we get

$$\begin{aligned}\operatorname{tr}(UTU^*) &= \operatorname{tr}(UT^{1/2}T^{1/2}U^*) = \operatorname{tr}((UT^{1/2})(UT^{1/2})^*) \\ &= \operatorname{tr}((UT^{1/2})^*(UT^{1/2})) = \operatorname{tr}(T^{1/2}U^*UT^{1/2}) = \operatorname{tr}(T).\end{aligned}$$

Given the orthonormal basis $(e_j)_{j\in J}$ through which the trace is defined, any other orthonormal basis for H is of the form $(Ue_j)_{j\in J}$ for some unitary $U \in \mathcal{B}(H)$. Hence

$$\sum_{j\in J}\langle T(Ue_j), Ue_j\rangle = \sum_{j\in J}\langle U^*TUe_j, e_j\rangle = \operatorname{tr}(U^*TU) = \operatorname{tr}(T).$$

This shows that the definition of the trace is independent of the chosen orthonormal basis. ■

The class of Hilbert–Schmidt operators were introduced in MAT4400, but we state the definition here:

Definition 5.1.4. An operator $T \in \mathcal{B}(H)$ is called *Hilbert–Schmidt* if

$$\operatorname{tr}(T^*T) = \sum_{j\in J}\langle T^*Te_j, e_j\rangle = \sum_{j\in J}\|Te_j\|^2 < \infty.$$

We denote by $\mathcal{HS}(H)$ the set of Hilbert–Schmidt operators on H.

We denote by $\mathcal{F}(H)$ the finite-rank operators on H, i.e., the operators $T \in \mathcal{B}(H)$ such that $\operatorname{Im}(T)$ is finite-dimensional. Recall that $\operatorname{Cl}(\mathcal{F}(H)) = \mathcal{K}(H)$, the compact operators on H.

Proposition 5.1.5. *The following hold for the Hilbert–Schmidt operators on a Hilbert space H:*

(a) $\mathcal{HS}(H)$ is a linear subspace of $\mathcal{K}(H)$ which contains $\mathcal{F}(H)$.

(b) $\mathcal{HS}(H)$ *is a Hilbert space with respect to the inner product*

$$\langle T, S\rangle = \operatorname{tr}(S^*T) \quad \textit{for } S, T \in \mathcal{HS}(H),$$

and the induced Hilbert–Schmidt norm is given by

$$\|T\|_2 = \operatorname{tr}(T^*T)^{1/2}.$$

(c) $\mathcal{HS}(H)$ *is a* $*$*-closed (two-sided) ideal in* $\mathcal{B}(H)$*, with*

$$\|ST\|_2 \le \|S\|_2\|T\|, \qquad \|TS\|_2 \le \|T\|\|S\|_2,$$
$$\|S^*\|_2 = \|S\|_2, \qquad \|S\| \le \|S\|_2$$

for $S \in \mathcal{HS}(H)$ *and* $T \in \mathcal{B}(H)$*.*

Proof. This was all done in MAT4400, so we leave it as an exercise to refresh on the proofs. ■

Definition 5.1.6. An operator $T \in \mathcal{B}(H)$ is called *trace class* if

$$\operatorname{tr}(|T|) < \infty.$$

We denote by $\mathcal{TC}(H)$ the set of trace class operators on H.

We need the following technical but important lemma:

Lemma 5.1.7. *Let* $T \in \mathcal{B}(H)$*. The following are equivalent:*

(a) $T \in \mathcal{TC}(H)$*.*

(b) There exist $R, S \in \mathcal{HS}(H)$ *such that* $T = RS$*.*

(c) $M'_T := \sup\Big\{ \sum_j |\langle Te_j, f_j\rangle| : (e_j)_j, (f_j)_j \textit{ are orthonormal sequences in } H\Big\} < \infty.$

Moreover, if $T \in \mathcal{TC}(H)$*, then*

$$\operatorname{tr}(|T|) = M'_T = M_T := \inf\{\|R\|_2\|S\|_2 : R, S \in \mathcal{HS}(H) \textit{ and } T = RS\}.$$

Proof. $(a) \Rightarrow (b)$: Let $T \in \mathcal{TC}(H)$. Let $T = U|T|$ be the polar decomposition of T. Set $R = U|T|^{1/2}$ and $S = |T|^{1/2}$. Then $S^*S = S^2 = |T|$, so $\operatorname{tr}(S^*S) = \operatorname{tr}(|T|) < \infty$. Thus $S \in \mathcal{HS}(H)$, so $R \in \mathcal{HS}(H)$ as well. We also get

$$\|R\|_2\|S\|_2 = \|U|T|^{1/2}\|_2\||T|^{1/2}\|_2 \le \|U\|\||T|^{1/2}\|_2\||T|^{1/2}\|_2 = \||T|^{1/2}\|_2^2 = \operatorname{tr}(|T|).$$

This shows that $M_T \le \operatorname{tr}(|T|)$.

$(b) \Rightarrow (c)$: Write $T = RS$ for some $R, S \in \mathcal{HS}(H)$. Let $(e_j)_j$ and $(f_j)_j$ be orthonormal sequences in H. Then

$$\begin{aligned}\sum_j |\langle Te_j, f_j\rangle| &= \sum_j |\langle Se_j, R^*f_j\rangle| \\ &\le \sum_j \|Se_j\|\|R^*f_j\|\end{aligned}$$

$$\leq \Big(\sum_j \|Se_j\|^2\Big)^{1/2}\Big(\sum_j \|R^*f_j\|^2\Big)^{1/2}$$
$$\leq \|S\|_2\|R^*\|_2 = \|S\|_2\|R\|_2.$$

Taking the supremum over all pairs of orthonormal sequences in H, we get $M'_T \leq \|R\|_2\|S\|_2$. Since R and S were general, it follows that $M'_T \leq M_T$.

$(c) \Rightarrow (a)$: Assume that $M'_T < \infty$. Let $(e_j)_j$ be an orthonormal basis for $\mathrm{Cl}(\mathrm{Im}\,|T|)$. Let $T = U|T|$ be the polar decomposition of T. Then $|T| = U^*T$ by Proposition 4.4.7. Since U is a partial isometry with initial space equal to $\mathrm{Cl}(\mathrm{Im}\,|T|)$ and final space equal to $\mathrm{Cl}(\mathrm{Im}\,T)$, it follows that $(Ue_j)_{j\in J}$ is an orthonormal basis for $\mathrm{Cl}(\mathrm{Im}\,T)$. Moreover, if we enlarge $(e_j)_j$ to an orthonormal basis $(f_k)_k$, then the new terms will form an orthonormal basis for $(\mathrm{Cl}(\mathrm{Im}\,|T|))^\perp = \mathrm{Ker}\,|T|$. Hence

$$\|T\|_1 = \sum_k \langle |T|f_k, f_k\rangle = \sum_j \langle |T|e_j, e_j\rangle = \sum_j \langle Te_j, Ue_j\rangle \leq M'_T.$$

In particular, $\mathrm{tr}(|T|) < \infty$.

We have now shown that (a), (b) and (c) are equivalent, and that if any of them hold, then $\mathrm{tr}(|T|) \leq M_T \leq M'_T \leq \mathrm{tr}(|T|)$. This finishes the proof. ∎

Theorem 5.1.8. *$\mathcal{TC}(H)$ is a $*$-closed ideal of $\mathcal{B}(H)$, and*

$$\|T\|_1 = \mathrm{tr}(|T|) \quad \text{for } T \in \mathcal{TC}(H)$$

defines a norm on $\mathcal{TC}(H)$ turning it into a Banach space. For $T \in \mathcal{TC}(H)$ and $S \in \mathcal{B}(H)$ we have

$$\|TS\|_1 \leq \|T\|_1\|S\|, \qquad \|ST\|_1 \leq \|S\|\|T\|_1,$$
$$\|T^*\|_1 = \|T\|_1, \qquad \|T\|_2 \leq \|T\|_1.$$

Finally, we have the following inclusions:

$$\mathcal{F}(H) \subseteq \mathcal{TC}(H) \subseteq \mathcal{HS}(H) \subseteq \mathcal{K}(H) \subseteq \mathcal{B}(H).$$

Proof. To begin with, we show that $\mathcal{TC}(H)$ is a linear subspace of $\mathcal{B}(H)$ and that $\|\cdot\|_1$ is a norm on $\mathcal{TC}(H)$. Let $T, S \in \mathcal{TC}(H)$ and $\lambda \in \mathbb{C}$. Let $(e_j)_j$ and $(f_j)_j$ be orthonormal sequences in H. Then

$$\sum_j |\langle (S+T)(e_j), f_j\rangle| \leq \sum_j |\langle Se_j, f_j\rangle| + \sum_j |\langle Te_j, f_j\rangle| \leq M'_S + M'_T.$$

Taking the supremum over all orthonormal sequences $(e_j)_j$ and $(f_j)_j$, we get by Lemma 5.1.7 that

$$\|S+T\|_1 = M'_{S+T} \leq M'_S + M'_T = \|S\|_1 + \|T\|_1.$$

In particular, $S + T \in \mathcal{TC}(H)$. It is easy to check that $\|\lambda T\|_1 = |\lambda|\|T\|_1$. Finally, if $\|T\|_1 = 0$, then

$$0 = \mathrm{tr}(|T|) = \sum_j \langle |T|e_j, e_j\rangle = \sum_j \||T|^{1/2}e_j\|^2.$$

Hence $\||T|^{1/2}e_j\| = 0$ for each j, so $|T|^{1/2} = 0$. This implies that $T = 0$. This proves that $\|\cdot\|_1$ is a norm on $\mathcal{TC}(H)$.

Next, we show that $\mathcal{TC}(H)$ is $*$-closed, and that $\|T^*\|_1 = \|T\|_1$ for all $T \in \mathcal{TC}(H)$: Let $T \in \mathcal{TC}(H)$, let (e_j) and $(f_j)_j$ be orthonormal sequences in H, and notice that

$$\sum_j |\langle T^*e_j, f_j\rangle| = \sum_j |\langle e_j, Tf_j\rangle| = \sum_j |\langle Tf_j, e_j\rangle| \leq M'_T < \infty.$$

It follows that $M'_{T^*} \leq M'_T$, so in particular $T^* \in \mathcal{TC}(H)$. Repeating the argument with $T = (T^*)^*$, we get $M'_T \leq M'_{T^*}$, so $\|T\|_1 = M'_T = M'_{T^*} = \|T^*\|_1$.

We now show that $\mathcal{TC}(H)$ is an ideal of $\mathcal{B}(H)$: Let $T \in \mathcal{TC}(H)$ and $S \in \mathcal{B}(H)$. By Lemma 5.1.7, we can write $T = RR'$ for $R, R' \in \mathcal{HS}(H)$. Since $\mathcal{HS}(H)$ is an ideal of $\mathcal{B}(H)$, $R'S, SR \in \mathcal{HS}(H)$, so $TS = R(R'S) \in \mathcal{TC}(H)$ and $ST = (SR)R' \in \mathcal{TC}(H)$. Moreover, by Proposition 5.1.5, we get

$$M_{TS} \leq \|R\|_2\|R'S\|_2 \leq \|R\|_2\|R'\|_2\|S\|.$$

Taking the infimum over all $R, R' \in \mathcal{HS}(H)$ such that $T = RR'$, we obtain

$$\|TS\|_1 = M_{TS} \leq M_T\|S\| = \|T\|_1\|S\|.$$

The proof that $\|ST\|_1 \leq \|S\|\|T\|_1$ is analogous.

By Lemma 5.1.7, every trace class operator can be written as a composition of two Hilbert–Schmidt operators. Since $\mathcal{HS}(H)$ is an ideal of $\mathcal{B}(H)$, we get $\mathcal{TC}(H) \subseteq \mathcal{HS}(H)$. In particular, trace class operators are compact. Let $T \in \mathcal{TC}(H)$. Then $|T| \in \mathcal{TC}(H)$, since we can write $|T| = U^*T$ where $T = U|T|$ is the polar decomposition of T. Hence $|T|$ is a compact, positive operator. By the spectral theorem for compact operators, we can find an orthonormal basis $(e_j)_{j\in J}$ for H consisting of eigenvectors for $|T|$: There are countably many nonnegative eigenvalues which can only accumulate at 0, so we can write $\lambda_1 \geq \lambda_2 \geq \cdots$ for them. Let f_n be the eigenvector from $(e_j)_{j\in J}$ corresponding to λ_n for each $n \in \mathbb{N}$. Then $(f_n)_{n\in\mathbb{N}}$ is an orthonormal basis for $(\operatorname{Ker}|T|)^\perp$. Now

$$\|T\|_1 = \sum_j \langle |T|e_j, e_j\rangle = \sum_n \langle |T|f_n, f_n\rangle = \sum_n \lambda_n\|f_n\|^2 = \sum_n \lambda_n$$

while

$$\|T\|_2^2 = \sum_j \|Te_j\|^2 = \sum_n \|Tf_n\|^2 = \sum_n \lambda_n^2.$$

Thus, since $(\sum_n \lambda_n^2)^{1/2} \leq \sum_n \lambda_n$, it follows that $\|T\|_2 \leq \|T\|_1$.

The completeness of $\mathcal{TC}(H)$ will follow from a later proposition. We leave the inclusion $\mathcal{F}(H) \subseteq \mathcal{TC}(H)$ as an exercise. ■

Let now $T \in \mathcal{TC}(H)$. For any orthonormal basis $(e_j)_j$ for H, the series $\sum_j \langle Te_j, e_j\rangle$ is absolutely convergent: Indeed, by Lemma 5.1.7, we have

$$\sum_j |\langle Te_j, e_j\rangle| \leq M'_T = \|T\|_1.$$

Note also that the series $\sum_j \langle Te_j, e_j\rangle$ is independent of the chosen orthonormal basis: We have already shown that this is the case when T is positive, and we can write more general operators as a linear combination of positive operators by first decomposing into real and imaginary parts and then decomposing into positive and negative parts. We can thus make the following definition:

Definition 5.1.9. Let $T \in \mathcal{TC}(H)$. Then we define the *trace* of T to be the (finite) number

$$\operatorname{tr}(T) = \sum_{j\in J} \langle Te_j, e_j\rangle$$

where $(e_j)_j$ is any orthonormal basis for H.

Note that for $T \in \mathcal{TC}(H)$ we have

$$|\operatorname{tr}(T)| = \Big|\sum_{j\in J} \langle Te_j, e_j\rangle\Big| \le \sum_{j\in J} |\langle Te_j, e_j\rangle| \le \|T\|_1.$$

Proposition 5.1.10. *Let $T \in \mathcal{TC}(H)$ and $S \in \mathcal{B}(H)$. Then*

$$\operatorname{tr}(TS) = \operatorname{tr}(ST).$$

Proof. Exercise. ■

We have now come to one of the main results regarding the trace class operators, which characterizes the Banach space $\mathcal{TC}(H)$ as the dual space of $\mathcal{K}(H)$. Before we present the theorem, recall that a finite rank operator can be written as a sum of rank one operators, and that every rank one operator $T \in \mathcal{B}(H)$ is of the form $T = \Theta_{y,z}$ for some $y, z \in H$, where

$$\Theta_{y,z}x = \langle x, z\rangle y \quad \text{for } x \in H.$$

We have $\|\Theta_{y,z}\| \le \|y\|\|z\|$ (check this if you have not seen it before).

Theorem 5.1.11. *For every $T \in \mathcal{TC}(H)$, the map $\omega_T \colon \mathcal{K}(H) \to \mathbb{C}$ given by*

$$\omega_T(S) = \operatorname{tr}(TS) \quad \textit{for } S \in \mathcal{K}(H)$$

defines a bounded, linear functional on $\mathcal{K}(H)$. Furthermore, the map $\omega \colon \mathcal{TC}(H) \to \mathcal{K}(H)^*$ given by

$$\omega(T) = \omega_T \quad \textit{for } T \in \mathcal{TC}(H)$$

is an isometric isomorphism.

Proof. Given $T \in \mathcal{TC}(H)$, linearity of the map ω_T is clear. Boundedness follows from

$$|\omega_T(S)| = |\operatorname{tr}(TS)| \le \|TS\|_1 \le \|T\|_1\|S\| \quad \text{for } S \in \mathcal{K}(H),$$

which also shows that $\|\omega_T\| \le \|T\|_1$.

Let $\phi \in \mathcal{K}(H)^*$. Define $L_\phi \colon H \times H \to \mathbb{C}$ by

$$L_\phi(x, y) = \phi(\Theta_{x,y}) \quad \text{for } x, y \in H.$$

Then L_ϕ is easily checked to be a sesquilinear form on H. Moreover, L_ϕ is bounded, as

$$|L_\phi(x,y)| \le \|\phi\|\|\Theta_{x,y}\| \le \|\phi\|\|x\|\|y\|$$

for all $x, y \in H$. By Proposition 4.6.6, there exists a unique $T \in \mathcal{B}(H)$ such that

$$\phi(\Theta_{x,y}) = \langle Tx, y\rangle \quad \text{for all } x, y \in H.$$

Assume y is a unit vector, and let $(e_j)_{j\in J}$ be an orthonormal basis for H in which y occurs. Then

$$\operatorname{tr}(T\Theta_{x,y}) = \sum_{j\in J}\langle T\Theta_{x,y}e_j, e_j\rangle = \sum_{j\in J}\langle T\langle e_j, y\rangle x, e_j\rangle = \sum_{j\in J}\langle Tx, e_j\rangle\langle e_j, y\rangle = \langle Tx, y\rangle = \phi(\Theta_{x,y}).$$

This formula is easily seen to extend to the case where y is not necessarily a unit vector as well. By linearity, we get that

$$\operatorname{tr}(TS) = \phi(S) \quad \text{for all finite rank operators } S \in \mathcal{B}(H).$$

We now show that T is trace class. Let $(e_j)_{j\in J}$ be an orthonormal basis. For each finite subset F of J, let P_F be the projection onto the finite-dimensional subspace spanned by $\{e_j : j \in F\}$. Let $T = U|T|$ be the polar decomposition of T, so that $|T| = U^*T$. Then P_FU^* is a finite rank operator, so $\operatorname{tr}(TP_FU^*) = \phi(P_FU^*)$ by what we have already proved. So for any finite $F \subseteq J$, we get

$$\begin{aligned}\sum_{j\in F}\langle |T|e_j, e_j\rangle &= \sum_{j\in J}\langle |T|P_Fe_j, e_j\rangle = \operatorname{tr}(|T|P_F)\\ &= \operatorname{tr}(U^*TP_F) = \operatorname{tr}(TP_FU^*) = \phi(P_FU^*)\\ &\le \|\phi\|\|P_F\|\|U^*\| = \|\phi\|.\end{aligned}$$

Hence, taking the supremum over all finite $F \subseteq J$, we get

$$\operatorname{tr}(|T|) = \sum_{j\in J}\langle |T|e_j, e_j\rangle = \lim_F \sum_{j\in F}\langle |T|e_j, e_j\rangle \le \|\phi\|.$$

This shows that $T \in \mathcal{TC}(H)$.

Let now $S \in \mathcal{K}(H)$. Then $S = \lim_{n\to\infty} F_n$ for a sequence of finite rank operators. Thus,

$$|\operatorname{tr}(TF_n) - \operatorname{tr}(TS)| = |\operatorname{tr}(T(F_n - S))| \le \|T(F_n - S)\|_1 \le \|T\|_1\|F_n - S\| \to 0.$$

This shows that

$$\phi(S) = \lim_{n\to\infty}\phi(F_n) = \lim_{n\to\infty}\operatorname{tr}(TF_n) = \operatorname{tr}(TS) = \omega_T(S).$$

Hence $\phi = \omega_T$. This shows surjectivity of ω, and we also get

$$\|T\|_1 \le \|\phi\| = \|\omega_T\| \le \|T\|_1.$$

Hence ω is isometric, which finishes the proof. ■

Theorem 5.1.12. *For every $T \in \mathcal{B}(H)$, the map $\psi_T \colon \mathcal{TC}(H) \to \mathbb{C}$ given by*

$$\psi_T(S) = \operatorname{tr}(TS) \quad \text{for } S \in \mathcal{TC}(H)$$

defines a bounded, linear functional on $\mathcal{TC}(H)$. Furthermore, the map $\psi \colon \mathcal{B}(H) \to \mathcal{TC}(H)^*$ given by

$$\psi(T) = \psi_T$$

is an isometric isomorphism.

Proof. Exercise. ■

5.2 Fredholm operators

Let H be a Hilbert space, let $y \in H$ and let $T \in \mathcal{B}(H)$. Suppose we wish to study the solutions of the equation $Tx = y$.

Uniqueness of solutions of this equation is, roughly speaking, related to the injectivity of T. When T is injective, i.e., when $\operatorname{Ker}(T) = \{0\}$, $Tx = y$ has at most one solution. The bigger the kernel is, the "less unique" solutions to the equation become.

On the other hand, existence of solutions is related to the surjectivity of T. When T is surjective, i.e., when $\operatorname{Im}(T) = H$, then solutions always exist. Letting

$$\operatorname{Coker}(T) := H/\operatorname{Im}(T),$$

we see that the bigger $\operatorname{Coker}(T)$ is, the "less likely" the equation is to have a solution.

There exists a unique solution if and only if T is invertible, if and only if $\operatorname{Ker}(T)$ and $\operatorname{Coker}(T)$ are both trivial. Considering operators for which these two spaces are finite-dimensional instead of trivial, we arrive at the definition of a Fredholm operator:

Definition 5.2.1. Let H be a Hilbert space. An operator $T \in \mathcal{B}(H)$ is called *Fredholm* if the vector spaces $\operatorname{Ker}(T)$ and $\operatorname{Coker}(T)$ are finite-dimensional. Moreover, the *index* of a Fredholm operator $T \in \mathcal{B}(H)$ is the integer

$$i(T) = \dim \operatorname{Ker}(T) - \dim \operatorname{Coker}(T).$$

Example 5.2.2. Let $T \in \mathcal{B}(H)$ be an invertible operator. Then as we have seen, T is Fredholm, and

$$i(T) = \dim \operatorname{Ker}(T) - \dim \operatorname{Coker}(T) = 0.$$

In particular, the identity operator I is always Fredholm, with $i(I) = 0$. If $S \in \mathcal{B}(H)$ is a Fredholm operator, then $\operatorname{Ker}(S) = \operatorname{Ker}(ST)$ and $\operatorname{Im}(S) = \operatorname{Im}(ST)$, so ST is Fredholm with index $i(ST) = i(T)$.

Example 5.2.3. Let H be a finite-dimensional Hilbert space and let $T \in \mathcal{B}(H)$. Then T is obviously Fredholm. Moreover, the rank-nullity theorem gives

$$i(T) = \dim \operatorname{Ker}(T) - \dim \operatorname{Coker}(T) = \dim \operatorname{Ker}(T) - (\dim H - \dim \operatorname{Im}(T)) = 0.$$

Proposition 5.2.4. *Let $T \in \mathcal{B}(H)$ be a Fredholm operator. Then $\operatorname{Im}(T)$ is closed in H.*

Proof. First, we prove the following claim: If M is a subspace of H such that H/M is finite-dimensional, then there exists a closed subspace N of H such that $M \cap N = \{0\}$ and $M + N = H$. Indeed, let $\{e_1 + M, \ldots, e_n + M\}$ be a basis for H/M. Set $N = \operatorname{span}\{e_1, \ldots, e_n\}$. Then N is finite-dimensional, hence closed. If $x \in M \cap N$, then we can find $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that $x = \sum_{i=1}^n \lambda_i e_i$. But then

$$0 + M = x + M = \sum_{i=1}^n \lambda_i (e_i + M),$$

so by linear independence we get $\lambda_i = 0$ for $1 \leq i \leq n$ Hence $x = 0$. Also, if $x \in H$, then we can find $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$x + M = \sum_{i=1}^n \lambda_i (e_i + M).$$

Setting $y = \sum_{i=1}^n \lambda_i e_i \in N$, we get $x - y \in M$ and $x = (x - y) + y \in M + N$. This finishes the proof of the claim.

Now let N be such a subspace for $M = \operatorname{Im}(T)$. We define a map $T' \colon (H/\operatorname{Ker}(T)) \oplus N \to H$ by

$$T'(x + \operatorname{Ker}(T), y) = T(x) + y \quad \text{for } (x, y) \in H \times N.$$

This is well-defined due to the linearity of T. Moreover, T' is bounded, linear and surjective since $\operatorname{Im}(T) + N = H$. and It is also injective since if $T(x) + y = 0$ for $x \in H$ and $y \in N$, then $T(x) = y = 0$ since $M \cap N = \{0\}$, so $(x + \operatorname{Ker}(T), y) = (\operatorname{Ker}(T), 0)$. By the open mapping theorem, T' is a homeomorphism. Hence

$$\operatorname{Im}(T) = T'((H/\operatorname{Ker}(T)) \times \{0\})$$

is closed since $(H/\operatorname{Ker}(T)) \times \{0\}$ is closed. ■

Corollary 5.2.5. *For an operator $T \in \mathcal{B}(H)$, the following are equivalent:*

(a) T is Fredholm.

(b) $\operatorname{Im}(T)$ is closed, and $\operatorname{Ker}(T)$ and $\operatorname{Ker}(T^*)$ are finite-dimensional.

Moreover, for T Fredholm, we have that $\dim \operatorname{Coker}(T) = \dim \operatorname{Ker}(T^*)$, so

$$i(T) = \dim \operatorname{Ker}(T) - \dim \operatorname{Ker}(T^*).$$

Proof. Note that for a closed subspace M of H, the restriction of the quotient map $H \to H/M$ to $M^\perp$ is a bijection, hence an invertible linear operator by the open mapping theorem (since H/M is a Banach space). In particular, if $T \in \mathcal{B}(H)$ has closed range, then $\operatorname{Im}(T) = \operatorname{Cl}(\operatorname{Im}(T)) = \operatorname{Ker}(T^*)^\perp$, so $H/\operatorname{Im}(T) \cong \operatorname{Ker}(T^*)$. From this $(b) \Rightarrow (a)$ immediately follows, and $(a) \Rightarrow (b)$ follows as well when combined with Proposition 5.2.4. ■

Example 5.2.6. Let $H = \ell^2(\mathbb{N})$ and let $S \in \mathcal{B}(H)$ be the shift operator given by

$$S(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots) \quad \text{for } x = (x_1, x_2, \ldots) \in H.$$

Then S is an isometry, so $\mathrm{Im}(S)$ is closed and $\mathrm{Ker}(S) = \{0\}$. The adjoint S^* is given by

$$S^*(x_1, x_2, \ldots) = (x_2, x_3, \ldots),$$

so $\mathrm{Ker}\, S^* = \mathrm{span}\{(x, 0, 0, \ldots) : x \in \mathbb{C}\}$. This shows that S is a Fredholm operator, with

$$i(S) = \dim \mathrm{Ker}(S) - \dim \mathrm{Ker}(S^*) = 0 - 1 = -1.$$

Lemma 5.2.7. *Let $T \in \mathcal{B}(H)$ be a compact operator. If M is a closed subspace of H and $M \subseteq \mathrm{Im}(K)$, then M is finite-dimensional.*

Proof. Exercise. ■

The following characterization of Fredholm operators is extremely useful:

Theorem 5.2.8 (Atkinson's Theorem)**.** *Let $T \in \mathcal{B}(H)$. The following are equivalent:*

(a) T is a Fredholm operator.

(b) There exists $S \in \mathcal{B}(H)$ such that both $ST - I$ and $TS - I$ are compact operators.

Proof. $(a) \Rightarrow (b)$: Suppose T is Fredholm. We claim that the operator $R\colon \mathrm{Ker}(T)^\perp \to \mathrm{Im}(T)$ given by $R(x) = T(x)$ is a bijection: It is certainly injective, since if $R(x) = 0$ for some $x \in \mathrm{Ker}(T)^\perp$, then $x \in \mathrm{Ker}(T) \cap \mathrm{Ker}(T)^\perp = \{0\}$. We also have

$$T(\mathrm{Ker}(T)^\perp) = T(\mathrm{Ker}(T) \oplus \mathrm{Ker}(T)^\perp) = T(H),$$

which shows that R is surjective.

Since $\mathrm{Im}(T)$ is closed, the open mapping implies that R has a bounded inverse $S\colon \mathrm{Im}(T) \to \mathrm{Ker}(T)^\perp$. Again, since $\mathrm{Im}(T)$ is closed, we can extend S to the whole of H by setting $S(y) = 0$ for $y \in \mathrm{Im}(T)^\perp$. For $x \in \mathrm{Im}(T)$ and $y \in \mathrm{Im}(T)^\perp$ we now get

$$TS(x + y) = RS(x + y) = RS(x) + RS(y) = x + R(0) = x.$$

Hence TS is the orthogonal projection onto $\mathrm{Im}(T) = \mathrm{Ker}(T^*)^\perp$. Furthermore, if $x' \in \mathrm{Ker}(T)$ and $y' \in \mathrm{Ker}(T)^\perp$, then

$$ST(x' + y') = ST(x') + ST(y') = S(0) + SR(y') = y'.$$

Hence ST is the orthogonal projection onto $\mathrm{Ker}(T)^\perp$.

It now follows that $I - TS$ is the orthogonal projection onto $\mathrm{Ker}(T^*)$, while $I - ST$ is the orthogonal projection onto $\mathrm{Ker}(T)$. Since these are finite-dimensional spaces by Corollary 5.2.5, the orthogonal projections onto them are finite rank operators, hence compact.

$(b) \Rightarrow (a)$: Let $S \in \mathcal{B}(H)$ be such that $I - ST$ and $I - TS$ are compact operators. Setting $K = ST - I \in \mathcal{K}(H)$, we have $ST = I + K$. If $x \in \mathrm{Ker}(I + K)$ then $Kx = -x$, so $x \in \mathrm{Im}(K)$. Hence $\mathrm{Ker}(T) \subseteq \mathrm{Ker}(ST) = \mathrm{Ker}(I + K) \subseteq \mathrm{Im}(K)$. By Lemma 5.2.7, $\mathrm{Ker}(T)$ must be finite-dimensional. Since $S^*T^* - I = (TS - I)^*$ is compact, a similar argument shows that $\mathrm{Ker}(T^*)$ is finite-dimensional.

By Corollary 5.2.5, it remains to show that $\mathrm{Im}(T)$ is closed. Since $K = ST - I$ is compact, there exists a finite rank operator $F \in \mathcal{B}(H)$ such that $\|K - F\| < 1/2$. If $x \in \mathrm{Ker}(F)$, then

$$\|S\|\|Tx\| \geq \|STx\| = \|(I + K)x\|$$

$$\geq \|x\| - \|Kx\| = \|x\| - \|(K-F)x\|$$
$$\geq \|x\| - \frac{1}{2}\|x\| = \frac{1}{2}\|x\|.$$

This shows that $T(\mathrm{Ker}(F))$ is closed: Indeed, if $(x_n)_n$ is a sequence in $\mathrm{Ker}(F)$ such that $(Tx_n)_n \to y \in H$, then

$$\|x_m - x_n\| \leq 2\|S\|\|Tx_m - Tx_n\|$$

which shows that $(x_n)_n$ is Cauchy. Hence $(x_n)_n$ has a limit $x \in \mathrm{Ker}(F)$ since the latter is closed, and $(T(x_n))_n \to T(x) \in T(\mathrm{Ker}(F))$.

Now $\mathrm{Ker}(F)^\perp = \mathrm{Cl}(\mathrm{Im}(F^*)) = \mathrm{Im}(F^*)$, which is finite-dimensional since F^* is a finite rank operator. Hence $T(\mathrm{Ker}(F)^\perp)$ is finite-dimensional. So

$$\mathrm{Im}(T) = T(H) = T(\mathrm{Ker}(F) \oplus \mathrm{Ker}(F)^\perp) = T(\mathrm{Ker}(F)) + T(\mathrm{Im}(F^*))$$

is expressible as the sum of a closed subspace and a finite-dimensional subspace. Such subspaces are closed: If $M \subseteq H$ is closed and $N \subseteq H$ is finite-dimensional, let $\pi\colon M \to H/M$ be the quotient map. Then $\pi(N) \subseteq H/M$ is finite-dimensional, hence closed, so $M + N = \pi^{-1}(\pi(N))$ is closed. ■

Recall that the *Calkin algebra* associated to a Hilbert space H is the Banach algebra $\mathcal{C}(H) = \mathcal{B}(H)/\mathcal{K}(H)$. Since $\mathcal{K}(H)$ is $*$-closed, the involution on $\mathcal{B}(H)$ passes to the Calkin algebra. In fact, $\mathcal{C}(H)$ becomes a C*-algebra with respect to this involution, but we will not need this.

Let $\pi\colon \mathcal{B}(H) \to \mathcal{C}(H)$, $T \mapsto \pi(T) = T + \mathcal{K}(H)$, be the quotient map. Then π is an algebra homomorphism, and is $*$-preserving by definition. Due to Theorem 5.2.8, Fredholm operators have a nice characterization in terms of the Calkin algebra: An operator $T \in \mathcal{B}(H)$ is Fredholm precisely when $\pi(T) \in \mathcal{C}(H)$ is invertible. This observation has some important consequences:

Proposition 5.2.9. *Let* $\mathrm{Fred}(H)$ *denote the set of Fredholm operators on* H. *Then* $\mathrm{Fred}(H)$ *has the following properties:*

(a) *It is an open subset of* $\mathcal{B}(H)$*,*

(b) *it is* $*$*-closed*

(c) *it is closed under composition,*

(d) *it is invariant under compact perturbations, that is, if* $T \in \mathrm{Fred}(H)$ *and* $K \in \mathcal{K}(H)$*, then* $T + K \in \mathrm{Fred}(H)$.

Proof. (a): Since $\mathrm{Fred}(H) = \pi^{-1}(GL(\mathcal{C}(H)))$ and the invertible elements of a unital Banach algebra form an open set, it follows that $\mathrm{Fred}(H)$ is open.

(b): If $T \in \mathrm{Fred}(H)$, then $\pi(T) \in GL(\mathcal{C}(H))$, so $\pi(T^*) = \pi(T)^* \in GL(\mathcal{C}(H))$, which means that $T^* \in \mathrm{Fred}(H)$.

(c): If $T, S \in \mathrm{Fred}(H)$, then $\pi(T), \pi(S) \in GL(\mathcal{C}(H))$. Since $\pi(TS) = \pi(T)\pi(S)$, we get that $\pi(TS) \in GL(\mathcal{C}(H))$ as well, so $TS \in \mathrm{Fred}(H)$.

(d): If $T \in \mathrm{Fred}(H)$ and $K \in \mathcal{K}(H)$, then

$$\pi(T + K) = \pi(T) + \pi(K) = \pi(T) \in GL(\mathcal{K}(H)),$$

so $T + K \in \mathrm{Fred}(H)$. ■

Set $\mathrm{Fred}_n(H) = \{T \in \mathrm{Fred}(H) : i(T) = n\}$ for every $n \in \mathbb{Z}$.

Lemma 5.2.10. *If $F \in \mathcal{B}(H)$ is a finite rank operator, then $I + F \in \mathrm{Fred}_0(H)$.*

Proof. Let $M = \mathrm{Im}(F) + \mathrm{Im}(F^*)$, which is a finite-dimensional (hence closed) subspace of H. We claim that M is invariant under $I+F$: Indeed, if we consider $Fx+F^*y$ for some $x, y \in H$, then

$$(I + F)(Fx + F^*y) = Fx + F^2x + F^*y + FF^*y \in M.$$

Similarly M is invariant under $I + F^*$ as well. We also claim that $(I + F)|_{M^\perp} = I_{M^\perp}$: Indeed, since $\mathrm{Im}(F^*) \subseteq M$ we have $M^\perp \subseteq \mathrm{Im}(F^*)^\perp = \mathrm{Ker}(F)$. Thus, if $x \in M^\perp$, then $(I + F)x = x + Fx = x$. Consequently

$$\mathrm{Im}(I + F) = (I + F)(H) = (I + F)(M) + (I + F)(M^\perp) = (I + F)(M) + M^\perp.$$

Since $(I + F)(M)$ is finite-dimensional and $M^\perp$ is closed, this shows that $\mathrm{Im}(I + F)$ is closed by the same argument as in the proof of Theorem 5.2.8.

We also claim that $\mathrm{Ker}(I+F) \subseteq M$: letting $x = x_1+x_2 \in H$ where $x_1 \in M$ and $x_2 \in M^\perp$, the assumption $(I + F)x = 0$ then gives

$$0 = (I + F)x_1 + (I + F)x_2 = (I + F)x_1 + x_2.$$

Here $(I + F)x_1 \in M$ by the invariance and $x_2 \in M^\perp$, so we get $x_2 = 0$. Hence $x = x_1 \in M$. Similarly $\mathrm{Ker}(I + F^*) \subseteq M$ as well. This shows in particular that both $\mathrm{Ker}(I + F)$ and $\mathrm{Ker}(I + F^*) = \mathrm{Ker}((I + F)^*)$ are finite-dimensional, being subspaces of M. Thus, we have shown that $I + F$ is Fredholm.

Set $T = I + M$. We now consider the restriction $T|_M$, which we can view as a bounded, linear operator on M. Since $\mathrm{Ker}(T), \mathrm{Ker}(T^*) \subseteq M$, we also have that $\mathrm{Ker}(T|_M) = \mathrm{Ker}(T)$ and $\mathrm{Ker}(T^*|_M) = \mathrm{Ker}(T^*)$. Since M is finite-dimensional, it follows from Example 5.2.3 that

$$i(T) = \dim \mathrm{Ker}(T) - \dim \mathrm{Ker}(T^*) = \dim \mathrm{Ker}(T|_M) - \dim \mathrm{Ker}(T^*|_M) = i(T|_M) = 0.$$

■

Proposition 5.2.11. *Let $T \in \mathrm{Fred}(H)$ and $T \in \mathcal{K}(H)$. Then $T + K \in \mathrm{Fred}(H)$, and*

$$i(T + K) = i(T).$$

Proof. Assume first that $i(T) = 0$. Then $\dim \mathrm{Ker}(T) = \dim \mathrm{Ker}(T^*) < \infty$. Pick a partial isometry $V \in \mathcal{B}(H)$ with initial space $\mathrm{Ker}(T)$ and final space $\mathrm{Ker}(T^*)$. Then V has finite rank. We claim that $T + V$ is bijective:

To see that $T+V$ is injective, let $x \in H$ and suppose that $(T+V)x = 0$. Writing $x = y+z$ where $y \in \mathrm{Ker}(T)$ and $z \in \mathrm{Ker}(T)^\perp$, we get

$$0 = (T + V)(y + z) = Tz + Vy.$$

Here $Tz \in \mathrm{Im}(T)$ while $Vy \in \mathrm{Ker}(T^*) = \mathrm{Im}(T)^\perp$, so we get $Tz = Vy = 0$. But then $z \in \mathrm{Ker}(T) \cap \mathrm{Ker}(T)^\perp$, so $z = 0$, and $y = 0$ since $V|_{\mathrm{Ker}(T)}$ is an isometry. Hence $x = 0$.

Surjectivity follows from the following computation:

$$(T + V)(H) = T(\mathrm{Ker}(T)^\perp) + V(\mathrm{Ker}(T)) = T(H) + \mathrm{Ker}(T^*)$$

$$= \mathrm{Ker}(T^*)^\perp + \mathrm{Ker}(T^*) = H.$$

The open mapping theorem now implies that $T + V \in GL(\mathcal{B}(H))$. Since $GL(\mathcal{B}(H))$ is open in $\mathcal{B}(H)$, we can find $\delta > 0$ such that $T + V + W \in GL(\mathcal{B}(H))$ whenever $\|W\| < \delta$.

Pick a finite rank operator $F \in \mathcal{B}(H)$ such that $\|K - F\| < \delta$. Then $S := T + V + K - F \in GL(\mathcal{B}(H))$. Set $G = F - V$. Then G has finite rank, and

$$T + K = S + F - V = S + G = S(I + S^{-1}G).$$

Since $S^{-1}G$ is a finite rank operator, $I + S^{-1}G \in \mathrm{Fred}_0(H)$ by Lemma 5.2.10. Since S is invertible, it follows that $T + K = S(I + S^{-1}G)$ is Fredholm, with

$$i(T + K) = i(S(I + S^{-1}G)) = i(I + S^{-1}G) = 0.$$

The case of more general $i(T) \in \mathbb{Z}$ is left as an exercise. ■

Proposition 5.2.12. *Let $S, T \in \mathcal{B}(H)$ be Fredholm operators. Then*

$$i(ST) = i(S) + i(T).$$

Proof. Assume first that $i(T) = 0$. As in the proof of Proposition 5.2.11, we can find a partial isometry $V \in \mathcal{B}(H)$ with finite rank such that $T + V$ is invertible. Therefore

$$i(S) + i(T) = i(S) + 0 = i(S(T + V)) = i(ST + SV) = i(ST).$$

The case of general $i(T) \in \mathbb{Z}$ is left as an exercise. ■

In light of the above proposition, we can view the index as a group homomorphism $i \colon GL(\mathcal{C}(H)) \to \mathbb{Z}$.

5.3 Spectral theory for compact operators

We will apply the theory of Fredholm operators to describe the spectrum of a compact operator, and give a new proof of the spectral theorem for normal, compact operators.

Theorem 5.3.1. *Let $K \in \mathcal{B}(H)$ be a compact operator. Then the following hold:*

(a) $\mathrm{sp}(K)$ *is countable.*

(b) If $\lambda \in \mathrm{sp}(K)$ and $\lambda \neq 0$, then λ is an eigenvalue of K, $\overline{\lambda}$ is an eigenvalue of K^*, and

$$\dim \mathrm{Ker}(\lambda I - K) = \dim \mathrm{Ker}(\overline{\lambda} I - K^*) < \infty.$$

(c) If $\mathrm{sp}(K)$ *is countably infinite and* $\mathrm{sp}(K) \setminus \{0\} = \{\lambda_1, \lambda_2, \ldots\}$, *then* $\lim_{n\to\infty} \lambda_n = 0$.

Proof. (*b*): Let $\lambda \in \mathrm{sp}(K) \setminus \{0\}$. Then λI is invertible, so it follows from Proposition 5.2.11 that $\lambda I - K$ is a Fredholm operator with $i(\lambda I - K) = 0$. Thus,

$$d = \dim \mathrm{Ker}(\lambda I - K) = \dim \mathrm{Ker}(\overline{\lambda} I - K^*) < \infty.$$

We cannot have $d = 0$: This would imply that $\operatorname{Ker}(\lambda I - K) = 0$ so that $\lambda I - K$ is injective, but also $\operatorname{Im}(\lambda I - K) = \operatorname{Cl}\operatorname{Im}(\lambda I - K) = \operatorname{Ker}(\overline{\lambda} I - K^*)^\perp = H$, giving that $\lambda I - K$ is invertible. This contradicts the fact that $\lambda \in \operatorname{sp}(K)$, so we conclude that $d \geq 1$. But then λ is an eigenvalue for K, and $\overline{\lambda}$ is an eigenvalue for K^*, and the corresponding eigenspaces are finite-dimensional, with the same dimension.

(a) and (c): Suppose $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of distinct eigenvalues of K. We will show that $(\mu_n)_n$ must go to zero. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of corresponding eigenvectors. For each $n \in \mathbb{N}$, set

$$M_n = \operatorname{span}\{x_1, \ldots, x_n\}.$$

Then $M_n \subset M_{n+1}$ for each $n \in \mathbb{N}$, so we can pick a sequence of unit vectors $(y_n)_n$ in H such that $y_1 \in M_1$ and $y_{n+1} \in M_{n+1} \cap M_n^\perp$ for all $n \in \mathbb{N}$. By construction, $(y_n)_n$ is an orthonormal sequence. Since K is compact and $(y_n)_n$ is orthonormal, we must have $K(y_n) \to 0$ as $n \to \infty$. (This is a general result which is covered in MAT4400. Exercise?)

Let $n \in \mathbb{N}$. Since $y_n \in M_n$, we can find $c_1, \ldots, c_n \in \mathbb{C}$ such that $y_n = \sum_{i=1}^n c_i x_i$. Then

$$Ky_n = \sum_{i=1}^n c_i K x_i = \sum_{i=1}^n c_i \mu_i x_i.$$

Set $\phi_n = Ky_n - \mu_n y_n$. Then

$$\phi_n = \sum_{i=1}^n c_i \mu_i x_i - \mu_n \sum_{i=1}^n c_i x_i = \sum_{i=1}^n c_i(\mu_i - \mu_n) x_i = \sum_{i=1}^{n-1} c_i(\mu_i - \mu_n) x_i \in M_{n-1}.$$

Consequently, the Pythagorean identity gives

$$|\mu_n|^2 = |\mu_n|^2 \|y_n\|^2 \leq \|\mu_n y_n\|^2 + \|\phi_n\|^2 = \|\mu_n y_n + \phi_n\|^2 = \|K(y_n)\|^2.$$

Letting $n \to \infty$, we get $\mu_n \to 0$.

This shows that *if* $\operatorname{sp}(K)$ was countably infinite, then any enumeration of $\operatorname{sp}(K)$ goes to zero. But it also shows that $\operatorname{sp}(K)$ is countable: For any $k \in \mathbb{N}$, the above argument shows that there can only be finitely many eigenvalues λ of K with $|\lambda| \geq 1/k$. Consequently

$$\operatorname{sp}(K) \setminus \{0\} = \bigcup_{k\in\mathbb{N}} \{\lambda \in \operatorname{sp}(K) : |\lambda| \geq 1/k\}$$

is countable. ∎

Theorem 5.3.2. *Let $K \in \mathcal{B}(H)$ be a normal, compact operator. Then H admits an orthonormal basis consisting of eigenvectors of K.*

Proof. For each $\lambda \in \operatorname{sp}(K)$, let S_λ be an orthonormal basis for $\operatorname{Ker}(\lambda I - K)$. Since K is normal, eigenvectors corresponding to different eigenvalues are orthogonal. It follows that $S = \bigcup_{\lambda\in\operatorname{sp}(K)} S_\lambda$ is an orthonormal set.

Let $M = \operatorname{Cl}(\operatorname{span} S)$. Then M contains all eigenvectors of K. Since both K and K^* leave M invariant, both K and K^* must also leave $M^\perp$ invariant. We can therefore consider $K|_{M^\perp} \in \mathcal{B}(M^\perp)$. This is a normal, compact operator on $M^\perp$, but has no eigenvalues: Any eigenvector for $K|_{M^\perp}$ would be an eigenvector for K in $M^\perp$, which would then be zero since it also lies in M. By Theorem 5.3.1, any nonzero element in the spectrum $K|_{M^\perp}$ is an eigenvalue, so we are forced to conclude that $\mathrm{sp}(K|_{M^\perp}) = \{0\}$. Since $K|_{M^\perp}$ is normal, this implies that $K|_{M^\perp} = 0$, i.e., $M^\perp \subseteq \mathrm{Ker}(K)$. But $\mathrm{Ker}(K)$ (if nontrivial) is the eigenspace of K corresponding to $\lambda = 0$, so $\mathrm{Ker}(K) \subseteq M$. Hence $M^\perp = 0$, so $M = H$ and the proof is finished. ■