Functional Analysis

Notes for Pure Math 453

Kenneth R. Davidson
University of Waterloo

CONTENTS

.

CHAPTER 1

Set Theory and Topology

1.1. Orders on Sets

First we define some important properties of orders which we require.

1.1.1. DEFINITION. A *partial order* on a set X is a relation $\leq$ satisfying

(1) $x \leq x$ for $x \in X$ (reflexive)

(2) $x \leq y$ and $y \leq x$ implies that $x = y$ (antisymmetric)

(3) $x \leq y$ and $y \leq z$ implies that $x \leq z$ (transitive).

There need not be a relation between two points $x, y \in X$. We call $(X, \leq)$ a *poset*.

A poset $(X, \leq)$ is a *total order* if given $x, y \in X$, then either $x \leq y$ or $x \geq y$. A totally ordered set $(X, \leq)$ is *well ordered* if every non-empty subset of X has a smallest element. $Y \subset X$ is an *initial segment* of a well-ordered set X if $y < x$ for all $y \in Y$ and $x \in X \setminus Y$.

A poset is *upward directed* if for $x_1, x_2 \in X$, there is a $y \in X$ so that $x_1 \leq y$ and $x_2 \leq y$. A totally ordered subset C of a poset X is called a *chain*. A poset is *inductive* if for every chain $C \subset X$, there is a *upper bound* $y \in X$ for C so that $x \leq y$ for every $x \in C$.

1.1.2. EXAMPLES.

(1) Let X be a set, and let $\mathcal{P}(X)$ denote the collection of all subsets of X. Put a relation on $\mathcal{P}(X)$ by setting $A \leq B$ if $A \subseteq B$. It is easy to see that this is a partial order. Moreover it is upward directed because given $A, B \in \mathcal{P}(X)$, $A \cup B$ is an upper bound. Indeed this set is inductive because if $\mathcal{C} \subset \mathcal{P}(X)$ is a chain, then $B = \bigcup_{A \in \mathcal{C}} A$ is an upper bound for $\mathcal{C}$.

(2) The real line $\mathbb{R}$ is totally ordered but not well ordered. For example $(0, 1)$ has no least element.

(3) The natural numbers $\mathbb{N}$ is well ordered.

(4) If $(X, \leq)$ is well ordered and infinite, then X has a least element x_1. Also $X \setminus \{x_1\}$ has a least element x_2. Recursively, we define x_{n+1} to be the least element of $X \setminus \{x_1, \dots, x_n\}$. Then $N = \{x_n : n \in \mathbb{N}\}$ is an initial segment of X which is order isomorphic to $\mathbb{N}$. This can be continued if $N \neq X$. Set x_ω to be the minimal element of $X \setminus N$, and set $x_{\omega+1}$ to be the minimal element of $X \setminus \{N \cup \{x_\omega\}\}$, etc.

1.2. The Axiom of Choice

The Axiom of Choice (AC) seems very natural, but it has some surprising consequences. It is known to be independent of the usual axioms of set theory (which we do not study), meaning that one can safely assume that it is valid, but can equally well assume other axioms instead which contradict AC. Most mathematicians other than set theorists and logicians assume the Axiom of Choice because it has many important positive consequences.

1.2.1. DEFINITION.
Axiom of Choice: for each non-empty set X, there is a *choice function*
$c : \mathcal{P}(X) \setminus \{\varnothing\} \to X$ so that $c(A) \in A$ for $\varnothing \neq A \subset X$.
Well Ordering Principle: every set X has a well-order.
Zorn's Lemma: every inductive partial order $(P, \leq)$ has a maximal element.

All of these properties are useful. Zorn's Lemma may seem to be the least intuitive. But in fact, it is probably the most useful of the three. The following result shows that they are all equivalent.

1.2.2. THEOREM. *The following are equivalent:*

(1) *Axiom of Choice*

(2) *Well Ordering Principle*

(3) *Zorn's Lemma*

PROOF. (2) implies (1). Put a well-order $\leq$ on X. For each non-empty set A, define $c(A)$ to be the least element in A in this order.

(3) implies (2). Let X be a set. Let

$$\mathcal{P} = \{(F, <_F) : F \subset X,\ <_F \text{ is a well-order on } F\}.$$

Say that $(F, <_F) \leq (G, <_G)$ if $F \subseteq G$, $<_F = <_G|_{F\times F}$ and F is an initial segment of G. It is easy to check that this is a partial order.

We will show that $(\mathcal{P}, \leq)$ is inductive: suppose that $\mathcal{C} = \{(F_\lambda, <_{F_\lambda}) : \lambda \in \Lambda\}$ is a chain in $\mathcal{P}$. Define $G = \bigcup_{\lambda\in\Lambda} F_\lambda$ and $<_G = \bigcup_{\lambda\in\Lambda} <_{F_\lambda}$. By the latter, I mean that if $x, y \in G$, then there are $\lambda, \mu \in \Lambda$ so that $x \in F_\lambda$ and $y \in F_\mu$. Since $\mathcal{C}$ is totally ordered, either $F_\lambda \subset F_\mu$ or $F_\mu \subset F_\lambda$; say $x, y \in F_\lambda$. Then since F_λ is a well order, either $x <_{F_\lambda} y$, $x = y$ or $y <_{F_\lambda} x$. Moreover, if $F_\lambda \subset F_\gamma$, then since the orders $<_{F_\lambda}$ and $<_{F_\gamma}$ agree on F_λ, the order on x and y is independent of the choice of λ. So $(G, <_G)$ is a total order. Finally if A is a non-empty subset of G, then there is some λ so that $A \cap F_\lambda \neq \varnothing$. Let a be the least element of $A \cap F_\lambda$ in $<_{F_\lambda}$. Then because F_λ is an initial segment of G, the element a satisfies $a <_G b$ for all $b \in A \setminus F_\lambda$. Therefore a is the least element of A in $(G, <_G)$. So G is well ordered. It is now clear that $(G, <_G)$ is an upper bound for $\mathcal{C}$. Therefore $\mathcal{P}$ is inductive.

By Zorn's Lemma, $\mathcal{P}$ has a maximal element $(F, <_F)$. If $F \neq X$, let $x \in X \setminus F$. Define a well order on $G = F \cup \{x\}$ by $a <_G b$ if $b = x$ or if $a, b \in F$ and $a <_F b$. Then G is well ordered and F is an initial segment of G. Thus $(F, <_F) \lneqq (G, <_G)$, contradicting the maximality of $(F, <_F)$. Therefore $F = X$ and we have a well order on X.

(1) implies (3). Let $(\mathcal{P}, \leq)$ be an inductive partial order. Suppose that $\mathcal{P}$ has no maximal elements. Then for each $x \in \mathcal{P}$, the set $U_x = \{y \in \mathcal{P} : x < y\}$ is non-empty. By the Axiom of Choice, there is a choice function c on non-empty subsets of $\mathcal{P}$. Define $f(x) = c(U_x)$.

For each chain $\mathcal{C} \subset \mathcal{P}$, let $U_{\mathcal{C}}$ denote the non-empty set of all upper bounds for $\mathcal{C}$. Define a function $g(\mathcal{C}) = c(U_{\mathcal{C}})$. Then $h(\mathcal{C}) = f(g(\mathcal{C}))$ is a *strict* upper bound for $\mathcal{C}$. If $x \in \mathcal{C}$, set $I(\mathcal{C}, x) = \{y \in \mathcal{C} : y < x\}$.

We will say (for this proof only) that $A \subset \mathcal{P}$ is *conforming* if $(A, \leq)$ is well ordered and for every $x \in A$, $h(I(A, x)) = x$. In particular, if a_1 is the least element of A, then $I(A, a_1) = \varnothing$ and $a_1 = h(\varnothing)$. Also if a_2 is the least element of $A \setminus \{a_1\}$, then $I(A, a_2) = \{a_1\}$ and $a_2 = h(a_1)$.

We claim that if A, B are conforming subsets of $\mathcal{P}$, then either A is an initial segment of B or B is an initial segment of A. To this end, let $\mathcal{X}$ consist of all initial segments which are common to A and B. Then $C = \bigcup \mathcal{X}$ is the largest initial segment common to A and B. If $C \in \{A, B\}$, then we are done. Otherwise let a be the least element of $A \setminus C$ and let b be the least element of $B \setminus C$. Then $I(A, a) = C = I(B, b)$. Since A and B are conforming, $a = h(C) = b$. Therefore $C \cup \{a\}$ is a larger initial segment of both A and B, a contradiction. So either $C = A$ or $C = B$.

Now $D = \bigcup\{A : A \text{ is conforming}\}$ must be the largest conforming subset of $\mathcal{P}$. However $D \cup h(D)$ is a strictly larger conforming set! This contradiction shows that the assumption that $\mathcal{P}$ has no maximal elements must be false. So (3) holds. ■

1.3. Topological spaces

1.3.1. DEFINITION. A *topology* τ on a set X is a collection of subsets such that

(1) $\varnothing,\ X \in \tau$.

(2) If $\{U_\lambda : \lambda \in \Lambda\} \subset \tau$, then $\bigcup_{\lambda \in \Lambda} U_\lambda \in \tau$.

(3) If $U_1, \dots, U_n \in \tau$, then $\bigcap_{i=1}^n U_i \in \tau$.

The elements $U \in \tau$ are called *open sets*.

1.3.2. EXAMPLES.
(1) If (X, d) is a metric space, then U is open if for every $x \in U$, there is an $r > 0$ so that the open ball $b_r(x) \subset U$.

(2) If X is any set, the *discrete topology* has $\tau_d = \mathcal{P}(X)$, the collection of all subsets of X.

(3) If X is any set, the *trivial topology* has $\tau = \{\varnothing, X\}$.

(4) If $(X, \leq)$ is a totally ordered set, the *order topology* is generated by the intervals $(a, b) = \{x \in X : a < x < b\}$, $(-\infty, b) = \{x \in X : x < b\}$ and $(a, \infty) = \{x \in X : x > a\}$. The topology consists of arbitrary unions of such intervals.

(5) If (X, τ) is a topology and $Y \subset X$, the *induced topology* on Y is $\tau|_Y = \{U \cap Y : U \in \tau\}$.

1.3.3. DEFINITION. A set $F \subset X$ is *closed* if F^c is open. If $A \subset X$, the *closure* of A is $\overline{A} = \bigcap\{F : A \subset F,\ F \text{ closed}\}$. A point in $\overline{A}$ is called a *limit point* of A.

If $A \subset X$, then $a \in A$ is an *interior point* of A if there exists $U \in \tau$ with $a \in U \subset A$. If $A \subset X$, the *interior* of A is A^o or $\operatorname{int} A = \bigcup\{U \in \tau : U \subset A\}$.

If $x \in X$, a *neighbourhood* of x is a set N such that $x \in N^o$.

1.3.4. PROPOSITION.

(1) *Finite unions and arbitrary intersections of closed sets are closed.*

(2) $\overline{A}$ *is the smallest closed set containing* A.

(3) $x \in \overline{A}$ *if and only if every* $U \in \tau$ *with* $x \in U$ *has* $A \cap U \neq \varnothing$.

(4) $\overline{A} = A^{coc}$ *is the complement of the interior of* A^c.

PROOF. Since open sets are closed under arbitrary unions and finite intersections, the collection of closed sets is closed under arbitrary intersections and finite unions. Hence the intersection of all closed sets $F \supset A$ is closed, and is thus the smallest closed set containing A. Now $x \in \overline{A}$ if and only if $x \in F$ for every closed $F \supset A$ if and only if $x \notin U$ if U is open and disjoint from A. Finally

$$X \setminus \overline{A} = \bigcup\{U \in \tau : U \cap A = \varnothing\} = \bigcup\{U \in \tau : U \subset A^c\} = A^{co}. \qquad \blacksquare$$

1.3.5. DEFINITION. If σ and τ are two topologies on X, we say that σ is a *weaker topology* than τ, and τ is a *stronger topology* than σ, if $\sigma \subset \tau$.

1.3.6. PROPOSITION. *If* $S \subset \mathcal{P}(X)$*, then there is a weakest topology* τ *containing* S*. It consists of arbitrary unions of sets which are intersections of finitely many elements of* S.

PROOF. Clearly if $\tau \supset S$ is a topology, then it contains all intersections of finitely many elements of S, and arbitrary unions of these sets. The intersection of no sets is X by convention, and $\varnothing$ is the union of no sets, so they both belong to τ. This collection is clearly closed under arbitrary unions. To check that it is stable under intersection, observe that if $A_{\alpha,i}$ and $B_{\beta,j}$ are in S, then

$$\bigcup_{\alpha\in A} A_{\alpha,1} \cap \cdots \cap A_{\alpha,n_\alpha} \;\cap\; \bigcup_{\beta\in B} B_{\beta,1} \cap \cdots \cap B_{\beta,m_\beta}$$
$$= \bigcup_{\alpha\in A,\,\beta\in B} A_{\alpha,1} \cap \cdots \cap A_{\alpha,n_\alpha} \cap B_{\beta,1} \cap \cdots \cap B_{\beta,m_\beta}.$$

Hence this collection is a topology. By construction, this is the weakest topology containing S. ■

1.3.7. DEFINITION. Say that $S \subset \mathcal{P}(X)$ is a *base* for a topology τ if every open set $U \in \tau$ is the union of elements of S. Also S is a *subbase* for a topology τ if the collection of finite intersections of elements of S is a base for τ.

1.3.8. EXAMPLES.
(1) If (X, d) is a metric space, then $\{b_{1/n}(x) : x \in X,\ n \geq 1\}$ is a base for the topology.

(2) $\{(r, s) : r < s \in \mathbb{Q}\}$ is a base for the topology of $\mathbb{R}$.

(3) Let $C[0, 1]$ denote the space of continuous functions on $[0, 1]$. For each $x \in [0, 1]$, $a \in \mathbb{C}$ and $r > 0$, let $U(x, a, r) = \{f \in C[0, 1] : f(x) \in b_r(a)\}$. Let τ be the topology generated by these sets. This is the topology of pointwise convergence. An open neighbourhood of f must contain a set of the form

$$\{g \in C[0, 1] : |g(x_i) - f(x_i)| < r \text{ for } 1 \leq i \leq n\}$$

for $x_1, \ldots, x_n \in [0, 1]$ and $r > 0$.

1.3.9. DEFINITION. A set A is *dense* in X if $X = \overline{A}$. X is *separable* if it has a countable dense subset. X is *first countable* if for each $x \in X$, there is a countable family $\{U_i\} \subset \tau$ with $x \in U_i$ which forms a *countable base of neighbourhoods of* x; i.e., if $x \in V$ is open, then there is some i so that $U_i \subset V$. X is *second countable* if there is a countable family of open sets which is a base for τ.

1.3.10. EXAMPLES.
(1) If (X, d) is a metric space and $x \in X$, then $\{b_{1/n}(x) : n \geq 1\}$ is a countable base of neighbourhoods of x. If X is separable, and $\{x_i : i \geq 1\}$ is dense in X, then $\{b_{1/n}(x_i) : i \geq 1,\ n \geq 1\}$ is a base for τ. Indeed, suppose that $x \in U$ is open. Pick $r > 0$ so that $b_r(x) \subset U$ and x_i so that $d(x, x_i) < 1/n < r/2$. Then

$x \in b_{1/n}(x_i) \subset U$. So X is second countable. In particular, compact metric spaces are separable and so second countable.

(2) Consider the discrete topology τ_d on a set X. Since the topology is generated by $\{\{x\} : x \in X\}$, X is always first countable. However it is second countable if and only if X is countable if and only if X is separable.

1.3.11. Definition. A topological space is T_0 if $x \neq y \in X$, then there is an open set containing one of these points, but not the other. A topological space is T_1 if points are closed. A topological space is *Hausdorff* or T_2 if for all $x, \neq y \in X$, there are open sets $U \ni x$ and $V \ni y$ so that $U \cap V = \varnothing$.

1.3.12. Examples.
(1) Let $X = \{0, 1\}$. Let $\tau = \{\varnothing, \{0\}, X\}$. Then $\{1\}$ is closed, but $\{0\}$ is not, and $\overline{\{0\}} = X$. Also $\{0\}$ is an open neighbourhood of 0 disjoint from 1. This space is T_0 but not T_1.

(2) Let $X = [0, 1) \cup \{a, b\}$. Let the open sets in τ be $U \subset [0, 1)$ which are open in the usual metric on $[0, 1)$ together with sets $U \cup (r, 1) \cup \{a\}$, $U \cup (r, 1) \cup \{b\}$ and $U \cup (r, 1) \cup \{a, b\}$ for $r < 1$. Here the points $\{a\}$ and $\{b\}$ are closed because the complement is open. However if $a \in U$ and $b \in V$ are open sets, then $U \cap V \supset (r, 1)$ for some $r < 1$. That means that you cannot separate a and b from one another by open sets, so it is not Hausdorff.

(3) Metric spaces are Hausdorff because if $x \neq y \in (X, d)$, then $d(x, y) = r > 0$. So $b_{r/2}(x)$ and $b_{r/2}(y)$ are disjoint open sets.

1.4. Nets

Sequences are not sufficient for dealing with convergence in general topological spaces, including many that arise in normal contexts. The replacement is the notion of a net, which you can think of as a very wide and very long generalized sequence.

1.4.1. Definition. A *net* in X is an upward directed poset Λ with a function $j : \Lambda \to X$, say $x_\lambda = j(\lambda)$. We usually write the net as $(x_\lambda)_\Lambda$. A net $(x_\lambda)_\Lambda$ *converges to* x in (X, τ) if for every open set $U \ni x$, there is $\lambda_0 \in \Lambda$ so that $x_\lambda \in U$ for every $\lambda \geq \lambda_0$. We write $\lim_\Lambda x_\lambda = x$.

A *subnet* $(y_\gamma)_\Gamma$ of $(x_\lambda)_\Lambda$ is given by a cofinal function $\varphi : \Gamma \to \Lambda$ so that $y_\gamma = x_{\varphi(\gamma)}$, where we say that φ is *cofinal* if for all $\lambda \in \Lambda$, there is a $\gamma_0 \in \Gamma$ so that $\varphi(\gamma) \geq \lambda$ for all $\gamma \geq \gamma_0$. It is convenient if φ is *monotone*, meaning that $\gamma_1 \leq \gamma_2$ implies that $\varphi(\gamma_1) \leq \varphi(\gamma_2)$. But this is not necessary.

We present a detailed example to explain why nets are needed, and how to use them.

1.4.2. Example. Let $X = \mathbb{N}_0 \times \mathbb{N}_0$. Declare that $U \subset X$ is open if $(0,0) \notin U$; and that a set $U \ni (0,0)$ is open if $\{m : \pi_1^{-1}(m) \cap U \text{ is cofinite in } \mathbb{N}_0\}$ is cofinite in $\mathbb{N}_0$. It is easy to verify that this defines a topology.

(a) X is Hausdorff because $\{(m,n)\}$ is open if $m+n \geq 1$ and $\{(m,n)\}^c$ is an open neighbourhood of $(0,0)$.

(b) $(0,0) \in \overline{X \setminus \{(0,0)\}}$ because every open set $U \ni (0,0)$ intersects $X \setminus \{(0,0)\}$.

(c) However no sequence $x_k = (m_k, n_k)$ in $X \setminus \{(0,0)\}$ converges to $(0,0)$. There are two cases. If $\{m_k : k \geq 1\}$ is bounded, pick m_0 so that $m_k = m_0$ infinitely often. The set $U = \{(m,n) : m \neq m_0\} \cup \{(0,0)\}$ is an open neighbourhood of $(0,0)$, and the sequence is not eventually in U. Otherwise, there is a sequence $k_i \to \infty$ so that $m_{k_i} < m_{k_{i+1}}$ for $i \geq 1$. Then $U = X \setminus \{x_{k_i} : i \geq 1\}$ is an open neighbourhood of $(0,0)$, and the sequence is not eventually in U.

(d) There is a net in $X \setminus \{(0,0)\}$ converging to $(0,0)$. Let $\Lambda = \{U \in \tau : (0,0) \in U\}$ where $U \leq V$ if $U \supset V$. (We say that Λ is ordered by containment.) This is directed because $U, V \leq U \cap V$. Order $X \setminus \{(0,0)\}$ by

$$(0,1), (1,0), (0,2), (1,1), (2,0), (0,3), (1,2), (2,1), (3,0), \ldots.$$

Define x_U to be the least element in this list which belongs to U. (This avoids any issues with the Axiom of Choice.) Then (x_U) converges to $(0,0)$ because given an open neighbourhood $U \ni (0,0)$, we have $x_V \in V \subset U$ whenever $U \leq V$.

(e) The sequence $(0,1), (1,0), (0,2), (1,1), (2,0), (0,3), (1,2), (2,1), \ldots$ has a subnet converging to $(0,0)$. Let Λ be the net just constructed. Define $\varphi(U) = x_U$ considered as an element in this sequence. To see that this map is cofinal, let (m_0, n_0) be in this sequence, and set $N_0 = m_0 + n_0$. Let

$$U_0 = X \setminus \{(m,n) : 1 \leq m+n \leq N_0\}.$$

Then if $U_0 \leq U$, it follows that $x_U = (m,n)$ with $m+n > N_0$ and thus $\varphi(U)$ follows (m_0, n_0) in the sequence. Therefore this is a subnet of the sequence which converges to $(0,0)$.

1.4.3. Proposition. *Let $A \subset X$. Then $x \in \overline{A}$ if and only if there is a net $(a_\lambda)_\Lambda$ in A such that $\lim_\Lambda a_\lambda = x$.*

Proof. Suppose that $x \in \overline{A}$. By Proposition 1.3.4(3), every open neighbourhood U of x intersects A. Let $\mathcal{O}(x)$ be the open neighbourhoods of x ordered by containment. By the Axiom of Choice, we can pick a point $a_U \in A \cap U$ for each $U \in \mathcal{O}(x)$. The net $(a_U)_{\mathcal{O}(x)}$ converges to x by construction.

Conversely, suppose that $(a_\lambda)_\Lambda$ is a net in A such that $\lim_\Lambda a_\lambda = x$. Then for any neighbourhood U of x, there is a λ so that $a_\lambda \in U$. In particular, $A \cap U \neq \varnothing$. Hence by Proposition 1.3.4(3), $x \in \overline{A}$. ■

1.5. Continuity

1.5.1. DEFINITION. A function $f : (X, \tau) \to (Y, \sigma)$ between topological spaces is *continuous* if for all $V \subset Y$ open, the set $f^{-1}(V)$ is open in X. Say that f is a *homeomorphism* if f is a bijection such that both f and f^{-1} are continuous.

1.5.2. EXAMPLES.

(1) The identity maps $(X, \text{discrete}) \xrightarrow{\iota_1} (X, \tau) \xrightarrow{\iota_2} (X, \text{trivial})$ is a continuous bijection, however in both cases ι_j^{-1} will be discontinuous provided that τ satisfies $\{\varnothing, X\} \subsetneq \tau \subsetneq \mathcal{P}(X)$.

(2) A function $f : (X, \text{trivial}) \to \mathbb{R}$ is continuous only if it is constant, while every function $f : (X, \text{discrete}) \to \mathbb{R}$ is continuous. On the other hand, a function $f : \mathbb{R} \to (X, \text{discrete})$ is continuous only if it is constant, while every function $f : \mathbb{R} \to (X, \text{trivial})$ is continuous.

(3) $f : (-1, 1) \to \mathbb{R}$ by $f(x) = \tan \frac{\pi x}{2}$ is a homeomorphism.

(4) Consider Example 1.3.12(1): $X = \{0, 1\}$ and $\tau = \big\{\varnothing, X, \{0\}\big\}$. If $f : X \to \mathbb{R}$ is continuous, then $f^{-1}(b_r(f(1)))$ is open and contains 1, so $f^{-1}(b_r(f(1)) = X$. Therefore f is constant.

(5) Consider Example 1.3.12(2): $X = [0, 1) \cup \{a, b\}$. If $a \in U$ and $b \in V$ are open sets, then $U \cap V \supset (r, 1)$ for some $r < 1$. Hence if $f : X \to \mathbb{R}$ is continuous, then

$$f^{-1}\big(b_\varepsilon(f(a))\big) \cap f^{-1}\big(b_\varepsilon(f(b))\big) \supset (r, 1) \quad \text{for some } r < 1.$$

This can only happen if $f(a) = f(b)$.

1.5.3. THEOREM. *Let $f : (X, \tau) \to (Y, \sigma)$. Then f is continuous if and only if whenever $(x_\lambda)_\Lambda$ is a net in X converging to x, it follows that $f(x) = \lim_\Lambda f(x_\lambda)$.*

PROOF. Suppose that f is continuous, and let $(x_\lambda)_\Lambda$ be a net in X converging to x. Let V be an open neighbourhood of $f(x)$. Then $U = f^{-1}(V)$ is an open neighbourhood of x. By convergence, there is a $\lambda_0 \in \Lambda$ so that $x_\lambda \in U$ for all $\lambda \geq \lambda_0$. Hence $f(x_\lambda) \in f(U) \subset V$ for all $\lambda \geq \lambda_0$. That means that $f(x) = \lim_\Lambda f(x_\lambda)$.

Conversely, suppose f is not continuous. Thus there is an open set $V \subset Y$ such that $U = f^{-1}(V)$ is not open. Then $\overline{U^c} \cap U$ contains a point x. By Proposition 1.4.3, there is a net $(x_\lambda)_\Lambda$ in U^c with limit x. Therefore $f(x_\lambda) \in f(U^c) \subset V^c$.

Since V^c is closed, any limit point of this net must remain in V^c by Proposition 1.4.3. Therefore it cannot converge to $f(x)$ which lies in V. So $f(x) \neq \lim_\Lambda f(x_\lambda)$. ■

1.5.4. DEFINITION. Let $C^b(X)$ and $C^b_{\mathbb{R}}(X)$ or $C^b(X, \mathbb{R})$ denote the normed vector space of bounded continuous functions from X into $\mathbb{C}$ and $\mathbb{R}$, respectively, with norm $\|f\|_\infty = \sup_X |f(x)|$. Similarly, $C(X)$ and $C_{\mathbb{R}}(X)$ or $C(X, \mathbb{R})$ denote the vector space of continuous functions from X into $\mathbb{C}$ and $\mathbb{R}$, respectively.

1.5.5. PROPOSITION. *If $C^b(X)$ separates points of X, i.e., for $x \neq y$ in X, there is a continuous function $f \in C^b(X)$ so that $f(x) \neq f(y)$, then X is Hausdorff.*

PROOF. If $f(x) = \alpha$ and $f(y) = \beta$ and $r = |\alpha - \beta|/2 > 0$, then $x \in U = f^{-1}(b_r(\alpha))$ and $y \in V = f^{-1}(b_r(\beta))$ and $U \cap V = \varnothing$. ■

Consider the Examples 1.5.2 (4) and (5) in light of this proposition.

Recall that $f_n \in C^b(X)$ *converge uniformly* to a function f if $\|f - f_n\|_\infty \to 0$. The following standard result for metric spaces extends easily.

1.5.6. PROPOSITION. *The uniform limit f of a sequence $f_n \in C^b(X)$ is continuous.*

PROOF. Let U be open in $\mathbb{C}$ and let $x \in f^{-1}(U)$. Then there is an $r > 0$ so that $b_r(f(x)) \subset U$. Choose n so large that $\|f - f_n\|_\infty < r/3$. Then $x \in V = f_n^{-1}(b_{r/3}(f_n(x)))$ is open. If $y \in V$, then $|f_n(y) - f_n(x)| < r/3$, so

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| \\ &< \|f - f_n\|_\infty + \frac{r}{3} + \|f - f_n\|_\infty < r. \end{aligned}$$

Hence $f(y) \in b_r(f(x)) \subset U$. Thus $V \subset f^{-1}(U)$. So f is continuous. ■

The norm $\|f\|_\infty$ makes $C^b(X)$ into a normed vector space. In view of Proposition 1.5.5, the following is most interesting when X is Hausdorff.

1.5.7. THEOREM. *For any topological space, $C^b(X)$ is complete.*

PROOF. Let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $C^b(X)$. If $\varepsilon > 0$, there is an N so that if $N \leq m < n$, then $\|f_n - f_m\|_\infty < \varepsilon$. In particular, for $x \in X$, the sequence $\big(f_n(x)\big)_{n\geq 1}$ is Cauchy in $\mathbb{C}$. So we may define $f(x) = \lim_{n\to\infty} f_n(x)$ pointwise. However for $m \geq N$,

$$|f(x) - f_m(x)| = \lim_{n\to\infty} |f_n(x) - f_m(x)| \leq \varepsilon.$$

Hence $\|f - f_m\|_\infty \leq \varepsilon$. So convergence is uniform. By Proposition 1.5.6, f is continuous. Also $\|f\|_\infty = \lim_{n\to\infty} \|f_n\|_\infty < \infty$, and so f lies in $C^b(X)$. Therefore $C^b(X)$ is complete. ■

1.6. Compactness

1.6.1. DEFINITION. An *open cover* of a set $A \subset X$ is a collection of open sets $\{U_\lambda : \lambda \in \Lambda\}$ such that $A \subset \bigcup_\Lambda U_\lambda$. A set A is *compact* if every open cover has a finite subcover, i.e., a finite subset $U_{\lambda_1}, \dots, U_{\lambda_n}$ such that $A \subset \bigcup_{i=1}^n U_{\lambda_i}$.

1.6.2. EXAMPLE. In 1.3.12(1), the point $\{0\}$ is compact but not closed.

1.6.3. PROPOSITION. *If X is compact and $A \subset X$ is closed, then A is compact.*

If X is Hausdorff and $A \subset X$ is compact, then A is closed. Moreover, if $x \notin A$, there are disjoint open sets $U \supset A$ and $V \ni x$.

PROOF. If $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ is an open cover of A, then $\mathcal{U} \cup \{A^c\}$ is an open cover of X. By compactness, it has a finite subcover $U_{\lambda_1}, \dots, U_{\lambda_n}, A^c$. Hence $U_{\lambda_1}, \dots, U_{\lambda_n}$ covers A; whence A is compact.

Suppose that X is Hausdorff and $A \subset X$ is compact, and let $x \in A^c$. For each $a \in A$, there are open sets $a \in U_a$ and $x \in V_a$ so that $U_a \cap V_a = \varnothing$. Clearly $\{U_a : a \in A\}$ is an open cover of A. By compactness, there is a finite subcover $U_{a_1}, \dots, U_{a_n}$. Let $V = \bigcap_{i=1}^n V_{a_i}$. Then $x \in V$ is open, and

$$V \cap A \subset \bigcup_{i=1}^n V \cap U_{a_i} = \varnothing.$$

Hence $x \notin \overline{A}$. So A is closed. Moreover $A \subset U = \bigcup_{i=1}^n U_{a_i}$, and $U \cap V = \varnothing$. ■

1.6.4. DEFINITION. A family $\{A_\lambda : \lambda \in \Lambda\}$ of subsets of X has the *finite intersection property* (FIP) if whenever $\lambda_1, \dots, \lambda_n$ are finitely many elements of Λ, then $\bigcap_{i=1}^n A_{\lambda_i} \neq \varnothing$.

1.6.5. PROPOSITION. *Let X be a topological space. The following are equivalent:*

(1) *X is compact.*

(2) *every family $\mathcal{F} = \{A_\lambda : \lambda \in \Lambda\}$ of closed sets with FIP has non-empty intersection $\bigcap \mathcal{F} := \bigcap_\Lambda A_\lambda \neq \varnothing$.*

(3) *every net in X has a convergent subnet.*

PROOF. Suppose that X is compact and $\mathcal{F}$ has FIP. Define open sets $U_\lambda = A_\lambda^c$. If $\bigcap \mathcal{F} = \varnothing$, then $\bigcup_\Lambda U_\lambda = \big(\bigcap \mathcal{F}\big)^c = X$. So $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ is an open cover of X. By compactness, there is a finite subcover $U_{\lambda_1}, \dots, U_{\lambda_n}$. Hence $\bigcap_{i=1}^n A_{\lambda_i} = \big(\bigcup_{i=1}^n U_{\lambda_i}\big)^c = \varnothing$, contradicting FIP. Therefore $\bigcap \mathcal{F} \neq \varnothing$.

Conversely, suppose that $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ is an open cover of X. Define closed sets $A_\lambda = U_\lambda^c$. If there is no finite subcover, then $\bigcap_{i=1}^n A_{\lambda_i} = \big(\bigcup_{i=1}^n U_{\lambda_i}\big)^c \neq \varnothing$; and thus $\mathcal{F} = \{A_\lambda : \lambda \in \Lambda\}$ has FIP. But then $\bigcap \mathcal{F} \neq \varnothing$. Therefore $\bigcup_\Lambda U_\lambda = \big(\bigcap \mathcal{F}\big)^c \neq X$, contradicting the fact that $\mathcal{U}$ is an open cover. Hence $\mathcal{F}$ does not have FIP, so there is a finite set $A_{\lambda_1}, \dots, A_{\lambda_n}$ such that $\bigcap_{i=1}^n A_{\lambda_i} = \varnothing$. Hence $\bigcup_{i=1}^n U_{\lambda_i} = \big(\bigcap_{i=1}^n A_{\lambda_i}\big)^c = X$. Therefore X is compact.

Suppose that (3) holds. Consider a collection $\mathcal{F} = \{C_\alpha : \alpha \in A\}$ of closed sets with the FIP. Let $\Lambda = \{F \subset A : F \text{ is finite, non-empty}\}$ ordered by inclusion, i.e., $F \le G$ if $F \subset G$. This is an upward directed poset: $F_1, F_2 \le F_1 \cup F_2$. For each $F \in \Lambda$, use the Axiom of Choice to select a point $x_F \in \bigcap_{\alpha \in F} C_\alpha$. This is possible since the finite intersection is non-empty. Then $(x_F)_\Lambda$ is a net in X. Let $(y_\gamma)_\Gamma$ be a subnet with limit x; where $\varphi : \Gamma \to \Lambda$ and $y_\gamma = x_{\varphi(\gamma)}$. For any $\alpha \in A$, there is a $\gamma_\alpha \in \Gamma$ so that $\gamma \ge \gamma_\alpha$ implies that $\varphi(\gamma) \ge \{\alpha\}$. Hence $y_\gamma \in C_\alpha$ for all $\gamma \ge \gamma_\alpha$. Since C_α is closed, the limit point $x \in C_\alpha$. This holds for all $\alpha \in A$. Therefore $x \in \bigcap \mathcal{F}$. Thus (2) holds.

Conversely suppose that (2) holds. Let $(x_\lambda)_\Lambda$ be a net in X. For each $\lambda \in \Lambda$, define $C_\lambda = \overline{\{x_\mu : \mu \ge \lambda\}}$. Then $\mathcal{F} = \{C_\lambda : \lambda \in \Lambda\}$ is a collection of non-empty closed sets. It has FIP because if $\lambda_1, \dots, \lambda_n \in \Lambda$, the upward directed property ensures that there is some $\lambda_0 \in \Lambda$ so that $\lambda_i \le \lambda_0$ for $1 \le i \le n$. Hence $\bigcap_{i=1}^n C_{\lambda_i} \supset C_{\lambda_0} \neq \varnothing$. Therefore, there is a point $x \in \bigcap_\Lambda C_\lambda$.

Now we build a subnet with limit x. Let $\mathcal{O}(x)$ be the set of all open neighbourhoods of x. Let $\Gamma = \Lambda \times \mathcal{O}(x)$ with order $(\lambda, U) \le (\mu, V)$ if $\lambda \le \mu$ and $U \supset V$. Let $S_{\lambda,U} = \{\mu \in \Lambda : \mu \ge \lambda \text{ and } x_\mu \in U\}$. This set is non-empty because $x \in C_\lambda \cap U = \overline{\{x_\mu : \mu \ge \lambda\}} \cap U$; and thus by Proposition 1.4.3, $x_\mu \in U$ for some $\mu \ge \lambda$. Use the Axiom of Choice to select $\mu = \varphi(\lambda, U) \in S_{\lambda,U}$ for each $(\lambda, U) \in \Gamma$. The map $\varphi : \Gamma \to \Lambda$ is cofinal because if $\lambda_0 \in \Lambda$, then every $(\lambda, U) \ge (\lambda_0, X)$ will have $\varphi(\lambda, U) = \mu \ge \lambda \ge \lambda_0$. So $y_{\lambda,U} = x_{\varphi(\lambda,U)}$ defines a subnet $(y_{\lambda,U})_\Gamma$ of $(x_\lambda)_\Lambda$.

Finally we claim that $\lim_\Gamma y_{\lambda,U)} = x$. Indeed, let $U \in \mathcal{O}(x)$ be any open neighbourhood of x. Fix some $\lambda_0 \in \Lambda$. Whenever $(\lambda, V) \ge (\lambda_0, U)$, we have $y_{\lambda,V} \in V \subset U$. Thus this net converges to x. ∎

1.6.6. PROPOSITION. *If $f : (X, \tau) \to (Y, \sigma)$ is continuous and $A \subset X$ is compact, then $f(A)$ is compact.*

PROOF. Let $\{V_\lambda : \lambda \in \Lambda\}$ is an open cover of $f(A)$ in Y. Define $U_\lambda = f^{-1}(V_\lambda)$. These are open sets by continuity, and they cover A. Thus there is a finite subcover $U_{\lambda_1}, \dots, U_{\lambda_n}$. Then since $V_\lambda \supset f(U_\lambda)$, it follows that $V_{\lambda_1}, \dots, V_{\lambda_n}$ covers $f(A)$. Hence $f(A)$ is compact. ■

The following important consequence follows directly.

1.6.7. EXTREME VALUE THEOREM. *If (X, τ) is compact and $f \in C(X)$, then $|f|$ attains it maximum. In particular, $\|f\|_\infty < \infty$.*

1.6.8. COROLLARY. *If X is a compact topological space, then $\|\cdot\|_\infty$ is a norm on $C(X)$; and $C(X)$ is complete.*

1.7. Weak Topologies

1.7.1. DEFINITION. Given a set X, let $\mathcal{F} = \{f_\alpha : X \to (Y_\alpha, \tau_\alpha) \mid \alpha \in A\}$ be a family of functions. There is a weakest topology $\tau_{\mathcal{F}}$ in which each f_α is continuous called the *weak topology* determined by $\mathcal{F}$. This is the topology generated by $\{f_\alpha^{-1}(U) : U \in \tau_\alpha,\ \alpha \in \Lambda\}$. A base for this topology is given by

$$\{f_{\alpha_1}^{-1}(U_1) \cap \cdots \cap f_{\alpha_n}^{-1}(U_n) : \alpha_i \in A,\ U_i \in \tau_{\alpha_i},\ n \geq 1\}.$$

1.7.2. EXAMPLE. If (X_α, τ_α) are topological spaces for $\alpha \in A$, we define the *product space* to be $X = \prod_A X_\alpha = \{x = (x_\alpha) : x_\alpha \in X_\alpha\}$ with the weakest topology τ which makes the coordinate projections $\pi_\alpha : X \to X_\alpha$ by $\pi_\alpha(x) = x_\alpha$ continuous.

The sets $\pi_\alpha^{-1}(U) = \prod_{\beta \in A \setminus \{\alpha\}} X_\beta \times U$ for $U \in \tau_\alpha$ are open and form a subbase for the topology. The product topology τ consist of arbitrary unions of finite intersections of the subbase. So if $\alpha_1, \dots, \alpha_n \in \Lambda$ and $U_i \in \tau_{\alpha_i}$, then the sets of the form

$$U_1 \times \cdots \times U_n \times \prod_{\beta \in A \setminus \{\alpha_i, 1 \leq i \leq n\}} X_\beta$$

form a base for the topology.

If A is finite, this is a familiar construction in the metric space case. Indeed, if (X_i, d_i) are metric spaces for $1 \leq i \leq n$, then

$$D\big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) = \max\big\{d_i(x_i, y_i) : 1 \leq i \leq n\big\}$$

is a metric on the product, and the metric topology coincides with the product topology.

When A is infinite, it often requires the Axiom of Choice to be able to say that X is non-empty.

1.7.3. PROPOSITION. *Let $\mathcal{F}$ be a family of functions $f_\alpha : X \to (Y_\alpha, \tau_\alpha)$ for $\alpha \in A$. Then a net $(x_\lambda)_\Lambda$ converges to $x \in (X, \tau_\mathcal{F})$ if and only if $f_\alpha(x_\lambda) \to f_\alpha(x)$ for all $\alpha \in A$.*

PROOF. If $(x_\lambda)_\Lambda \to x$ in $(X, \tau_\mathcal{F})$, then since each f_α is continuous by definition, $f_\alpha(x_\lambda) \to f_\alpha(x)$ by Theorem 1.5.3.

Conversely, let U be an open neighbourhood of x. Then x has a basic open neighbourhood $x \in V = f_{\alpha_1}^{-1}(U_1) \cap \cdots \cap f_{\alpha_n}^{-1}(U_n) \subset U$ for some $\alpha_i \in A$ and $U_i \in \tau_{\alpha_i}$. For each $1 \le i \le n$, there is some $\lambda_i \in \Lambda$ so that $f_{\alpha_i}(x_\lambda) \in U_i$ for all $\lambda \ge \lambda_i$. Since Λ is upward directed, there is some $\lambda_0 \ge \lambda_i$ for $1 \le i \le n$. When $\lambda \ge \lambda_0$, $f_{\alpha_i}(x_\lambda) \in U_i$ for $1 \le i \le n$. Therefore $x_\lambda \in V \subset U$. It follows that $(x_\lambda)_\Lambda$ converges to x. ■

1.7.4. COROLLARY. *Let $\mathcal{F}$ be a family of functions $f_\alpha : X \to (Y_\alpha, \tau_\alpha)$ for $\alpha \in A$. Then $g : (Y, \tau) \to (X, \tau_\mathcal{F})$ is continuous if and only if $f_\alpha \circ g : Y \to Y_\alpha$ are continuous for all $\alpha \in A$.*

PROOF. The composition of continuous functions is continuous. So if g is continuous, so are $f_\alpha \circ g$ for all $\alpha \in A$. Conversely, suppose that a net $(y_\lambda)_\Lambda$ converges to $y \in Y$. By continuity, $f_\alpha(g(y_\lambda))$ converges to $f_\alpha(g(y))$ in Y_α for all $\alpha \in A$. By Proposition 1.7.3, $g(y_\lambda)$ converges to $g(y)$ in Y. Therefore g is continuous by Theorem 1.5.3. ■

1.7.5. PROPOSITION. *Let $\mathcal{F}$ be a family of functions $f_\alpha : X \to (Y_\alpha, \tau_\alpha)$ for $\alpha \in A$. If the Y_α are all Hausdorff and $\mathcal{F}$ separates points, then $(X, \tau_\mathcal{F})$ is Hausdorff.*

PROOF. Let $x \ne y \in X$. Since $\mathcal{F}$ separates points, there is some α so that $f_\alpha(x) \ne f_\alpha(y)$. Since Y_α is Hausdorff, there are disjoint open neighbourhoods $U \ni f_\alpha(x)$ and $V \ni f_\alpha(y)$. Therefore $f_\alpha^{-1}(U) \ni x$ and $f_\alpha^{-1}(V) \ni y$ are disjoint open sets in X. Thus X is Hausdorff. ■

1.7.6. COROLLARY. *If (X_α, τ_α) are Hausdorff spaces for $\alpha \in A$, then $X = \prod_{\alpha \in A} X_\alpha$ is Hausdorff in the product topology.*

1.7.7. TYCHONOFF'S THEOREM. *The product of compact spaces is compact.*

PROOF. Let $X = \prod_{\alpha \in A} X_\alpha$ where each X_α is compact. Suppose that there is an open cover with no finite subcover. Order

$$\mathcal{P} = \{\mathcal{U} : \text{open cover with no finite subcover}\}$$

by inclusion. If $\mathcal{C} = \{\mathcal{U}_\beta : \beta \in B\}$ is a chain in $\mathcal{P}$, let $\mathcal{U} = \bigcup_{\beta \in B} \mathcal{U}_\beta$. This is an open cover of X. If it has a finite subcover $U_1, \dots, U_n$, the chain property shows that there is some β so that $U_i \in \mathcal{U}_\beta$ for $1 \le i \le n$. Hence $\mathcal{U}_\beta$ contains a finite subcover, contrary to fact. Therefore $\mathcal{U} \in \mathcal{P}$ is an upper bound for $\mathcal{C}$. Therefore $\mathcal{P}$ is inductive. By Zorn's Lemma, there is a maximal open cover $\mathcal{U}_0$ with no finite subcover.

We claim the $\mathcal{U}_0$ has the following properties:

(1) If $U \in \mathcal{U}_0$ and $V \subset U$ is open, then $V \in \mathcal{U}_0$. This is because V can't help cover anything that U doesn't cover.

(2) If $U_1, U_2 \in \mathcal{U}_0$, then $U_1 \cup U_2 \in \mathcal{U}_0$. This is because $U_1 \cup U_2$ can't cover anything not covered by U_1 and U_2 together.

(3) If V_1, V_2 are open and $V_1 \cap V_2 \in \mathcal{U}_0$, then one of $V_i \in \mathcal{U}_0$. Indeed, if both $V_i \notin \mathcal{U}_0$, then $\{V_i, \mathcal{U}_0\}$ has a finite subcover, necessarily of the form $V_1 \cup U_1 \cup \dots \cup U_m = X = V_2 \cup U_{m+1} \cup \dots U_{m+n}$. However then $X = (V_1 \cap V_2) \cup U_1 \cup \dots \cup U_{m+n}$ is a finite subcover from $\mathcal{U}_0$, a contradiction.

For each $\alpha \in A$, let $\mathcal{Y}_\alpha = \{U \in \tau_\alpha : \pi_\alpha^{-1}(U) \in \mathcal{U}_0\}$ and let $Y_\alpha = \bigcup \mathcal{Y}_\alpha$. Then $Y_\alpha \neq X_\alpha$. For otherwise, $\mathcal{Y}_\alpha$ would be an open cover of X_α, and hence would have a finite subcover $U_1, \dots, U_n$. Then $\pi_\alpha^{-1}(U_1), \dots, \pi_\alpha^{-1}(U_n)$ would be a finite subcover from $\mathcal{U}_0$. Hence $X_\alpha \setminus Y_\alpha$ is non-empty for each $\alpha \in A$. By the Axiom of Choice, there is a point $x = (x_\alpha) \in X$ so that $x_\alpha \in X_\alpha \setminus Y_\alpha$ for each $\alpha \in A$. There must be a set $U \in \mathcal{U}_0$ with $x \in U$. Hence there is a basic open neighbourhood $x \in V = \pi_{\alpha_1}^{-1}(U_1) \cap \dots \cap \pi_{\alpha_n}^{-1}(U_n) \subset U$. Therefore, $\bigcap_{i=1}^n \pi_{\alpha_i}^{-1}(U_i) \in \mathcal{U}_0$ by property (1). Repeated application of property (3) now shows that there is some i so that $\pi_{\alpha_i}^{-1}(U_i) \in \mathcal{U}_0$. But this means that $x_{\alpha_i} \in Y_{\alpha_i}$, contrary to our construction of x. This contradiction shows that every open cover has a finite subcover. ■

We finish off this section by showing that Tychonoff's Theorem implies the Axiom of Choice. So it is in fact equivalent to Choice, since Choice was used to prove Tychonoff's Theorem.

1.7.8. THEOREM. *Tychonoff's Theorem is equivalent to the Axiom of Choice.*

PROOF. Beyond basic rules of set theory, we used the Axiom of Choice and the equivalent Zorn's Lemma to establish Tychonoff's Theorem.

Conversely, suppose that X_α for $\alpha \in A$ are non-empty sets. Define topological spaces $Y_\alpha = X_\alpha \dot\cup \{p_\alpha\}$, where we have added a distinguished point p_α with the topology $\tau_\alpha = \{\varnothing, \{p_\alpha\}, X_\alpha, Y_\alpha\}$. Clearly each Y_α is compact because every cover is finite. By Tychonoff's Theorem, $Y = \prod_{\alpha \in A} Y_\alpha$ is compact. Since X_α is closed, $C_\alpha = X_\alpha \times \prod_{\beta \neq \alpha} Y_\beta$ is closed in Y. A closed subset of a compact set is compact, so C_α is compact. The collection $\mathcal{F} = \{C_\alpha : \alpha \in A\}$ has FIP because if $\alpha_1, \dots, \alpha_n$ are given, we can choose $x_i \in X_{\alpha_i}$ for $1 \le i \le n$. Then $\bigcap_{i=1}^n C_{\alpha_i} \ni x = (x_1, \dots, x_n, p_\beta) : \beta \in A \setminus \{\alpha_1, \dots, \alpha_n\}$. By Proposition 1.6.5, the intersection $C = \bigcap_{\alpha \in A} C_\alpha$ is non-empty. Let $x = (x_\alpha)$ be a point in the intersection. Note that $x_\alpha \in X_\alpha$ for all $\alpha \in A$. Thus $c(\alpha) = x_\alpha$ is a choice function. Hence AC holds. ■

1.8. Compact Hausdorff Spaces*

1.8.1. DEFINITION. A topological space is *normal* if given disjoint closed sets A, B, there are disjoint open sets $U \supset A$ and $V \supset B$.

1.8.2. PROPOSITION. *Compact Hausdorff spaces are normal.*

PROOF. Let A, B be disjoint closed subsets of a compact Hausdorff space X. Then A and B are compact by Proposition 1.6.3. Moreover since X is Hausdorff, that same Proposition shows that for each point $x \in B$, there are disjoint open sets $U_x \supset A$ and $V_x \ni x$. The collection $\{V_x : x \in B\}$ is an open cover of B. Let $V_{x_1}, \dots, V_{x_n}$ be a finite subcover. Set $V = \bigcup_{i=1}^n V_{x_i}$ and $U = \bigcap_{i=1}^n U_{x_i}$. These are disjoint open sets with $A \subset U$ and $B \subset V$. ■

Now we prove that normal Hausdorff spaces have lots of continuous functions.

1.8.3. URYSOHN'S LEMMA. *Let X be a normal Hausdorff space, and let A and B be disjoint closed sets. Then there is a continuous function $f : X \to [0, 1]$ such that $f|_A = 0$ and $f|_B = 1$.*

PROOF. Normality implies the following property: if A is closed and W is open and $A \subset W$, then there is an open set U such that $A \subset U \subset \overline{U} \subset W$. To see this, take $B = W^c$. Use normality to find disjoint open sets $U \supset A$ and $V \supset B$. Then $\overline{U} \subset V^c \subset W$.

Start with $U_1 = B^c$. Find an open $U_{1/2}$ so that $A \subset U_{1/2} \subset \overline{U_{1/2}} \subset U_1$. Repeating this procedure recursively, we find open sets $U_{k/2^n}$ for $1 \le k \le 2^n$ and $n \ge 1$ so that

$$A \subset U_{k/2^n} \subset \overline{U_{k/2^n}} \subset U_{(k+1)/2^n} \quad \text{for} \quad 1 \le k < 2^n.$$

Let $D = \{k/2^n : 1 \le k \le 2^n,\ n \ge 1\}$. Define $f(x) = \inf\{r \in D : x \in U_r\}$ if $x \in U_1$ and $f|_B = 1$. Clearly $0 \le f \le 1$ and $f|_A = 0$.

<u>Claim:</u> f is continuous. Note that

$$f^{-1}\big([0, t)\big) = \bigcup_{r<t,\, r \in D} U_r \quad \text{is open for } t \in [0, 1].$$

Also for $0 \le t < 1$, since $t < r < s$ for $r, s \in D$ implies that $\overline{U_r} \subset U_s$,

$$f^{-1}\big([0, t]\big) = \bigcap_{r>t,\, r \in D} f^{-1}\big([0, r)\big) = \bigcap_{r>t,\, r \in D} U_r = \bigcap_{r>t,\, r \in D} \overline{U_r}.$$

This is closed, and therefore $f^{-1}\big((t,1]\big) = \big(\bigcap_{r>t,\, r\in D} \overline{U_r}\big)^c$ is open. Hence, $f^{-1}\big((s,t)\big)$ is open for $s < t$, and so f is continuous. ■

1.8.4. REMARK. If (X,d) is a metric space and A and B are disjoint closed sets, define

$$f(x) = \frac{d(x,A)}{d(x,A) + d(x,B)}.$$

This satisfies the conclusion of Urysohn's Lemma.

1.8.5. COROLLARY. *If X is a compact Hausdorff space, $C(X)$ separates points.*

Urysohn's Lemma implies the following significant strengthening.

1.8.6. TIETZE'S EXTENSION THEOREM. *Let X be a normal Hausdorff space, and let $A \subset X$ be a closed set. If $f : A \to [a,b]$ is continuous, there is a continuous function $F : X \to [a,b]$ such that $F|_A = f$.*

PROOF. After scaling, we may assume that the range is $[-1,1]$. Let $A_1 = f^{-1}([-1,-\frac{1}{3}])$ and $B_1 = f^{-1}([\frac{1}{3},1])$. By Urysohn's Lemma, there is a function $g_1 : X \to [-\frac{1}{3},\frac{1}{3}]$ so that $g_1|_{A_1} = -\frac{1}{3}$ and $g_1|_{B_1} = \frac{1}{3}$. Then $f_1 = f - g_1|_A$ has range in $[-\frac{2}{3},\frac{2}{3}]$. Repeat the process, setting $A_2 = f_1^{-1}([-\frac{2}{3},-\frac{2}{9}])$ and $B_2 = f_1^{-1}([\frac{2}{9},\frac{2}{3}])$, and finding $g_2 : X \to [-\frac{2}{9},\frac{2}{9}]$ with $g_2|_{A_2} = -\frac{2}{9}$ and $g_2|_{B_2} = \frac{2}{9}$.

Then $f_2 = f_1 - g_2|_A$ has range in $[-\big(\frac{2}{3}\big)^2, \big(\frac{2}{3}\big)^2]$. Recursively we obtain functions $g_n : X \to [-2\cdot 3^{-n}, 2\cdot 3^{-n}]$ so that $f_n = f - \sum_{i=1}^n g_n|_A$ has range in $[-\big(\frac{2}{3}\big)^n, \big(\frac{2}{3}\big)^n]$. Let $g = \sum_{n\ge 1} g_n$. Then $g|_A = f$ and

$$\|g\|_\infty \le \sum_{n\ge 1} \|g_n\|_\infty = \sum_{n\ge 1} 2\cdot 3^{-n} = 1.$$ ■

Exercises for Chapter 1

1. (a) Show that compact subsets of Hausdorff spaces are closed.
(b) Let $h : X \to Y$ be a continuous bijection of X onto Y. Suppose that X is compact and Y is Hausdorff. Prove that X and Y are homeomorphic.

2. In a topological space X, say that a net $(x_\lambda)_{\lambda\in\Lambda}$ has x as a *cluster point* if for each $\lambda_0 \in \Lambda$ and each open neighbourhood $U \ni x$, there is some $\lambda \ge \lambda_0$ so that $x_\lambda \in U$. Prove that x is a cluster point of this net if and only if there is a subnet with limit x. Be explicit about use of the Axiom of Choice.

3. Let V be the vector space of complex valued functions on $\mathbb{R}$. Put a topology τ on V with a subbase given by $U_{t,a,r} = \{f \in V : |f(t) - a| < r\}$ for $t \in \mathbb{R}$, $a \in \mathbb{C}$, $r > 0$.

(a) Show that a net $(f_\lambda)_{\lambda \in \Lambda}$ in V converges if and only if $\lim f_\lambda(t)$ exists for each $t \in \mathbb{R}$. In particular, $\varepsilon_t(f) := f(t)$ is a continuous function on V for each $t \in \mathbb{R}$.

(b) Let $E = \{f \in V : f(t) = 2$ except for finitely many $t \in \mathbb{R}\}$. Construct a net in E with limit 0.

(c) Show that no sequence of points in E can converge to 0.

4. Let $\mathbb{R}^{\mathbb{R}}$, the set of real functions on $\mathbb{R}$, have the product topology.

(a) Show that a net $(f_\lambda)_{\lambda \in \Lambda}$ converges if and only if $\lim f_\lambda(x)$ exists for each $x \in \mathbb{R}$.

(b) Let $E = \{f \in \mathbb{R}^{\mathbb{R}} : f(x) = 2$ except finitely often$\}$. Show that the zero function z is in the closure of E.

(c) Construct a net in E with limit z.

(d) Show that no sequence of points in E can converge to z.

5. Consider $\mathbb{R}$ with its usual metric topology τ, which is also the order topology on $\mathbb{R}$. Let Y be a subset of $\mathbb{R}$. Consider the topology ρ on Y induced by the topology on $\mathbb{R}$, and let σ be the order topology on Y given by the induced order.

(a) What relationship always holds between ρ and σ?

(b) Find a subset Y where these two topologies are different.

(c) Find reasonable conditions on Y which ensure that these two topologies agree. **Remark.** Try to find necessary and sufficient conditions. Reasonable means that your conditions should handle the cases where Y is open, closed, $\mathbb{Q}$ and your example in part (b).

6. (a) Prove that there is an uncountable well-ordered set Ω_0 such that the initial segments $I(a) = \{x \in \Omega_0 : x < a\}$ are countable for every $a \in \Omega_0$. Endow Ω_0 with the order topology.

HINT: well-order $\mathbb{R}$ and select an appropriate initial segment.

(b) Show that every increasing sequence in Ω_0 converges.

(c) Show that every sequence in Ω_0 has a convergent subsequence (i.e. Ω_0 is *sequentially compact*).

(d) Show that Ω_0 is not compact.

(e) Show that every continuous function $f : \Omega_0 \to \mathbb{R}$ is eventually constant.

HINT: show there is an $a_k \in \Omega_0$ so that if $b > a_k$, then $|f(b) - f(a_k)| < \frac{1}{k}$.

7. The *Moore plane* is $\Gamma = \{(x, y) : x \in \mathbb{R}, y \geq 0\}$ where for each (x, y) with $y > 0$, the usual open balls $B_r((x, y))$ of radius $r \leq y$ are open; and for $(x, 0) \in R = \{(x, 0) : x \in \mathbb{R}\}$, the sets $\{(x, 0)\} \cup B_r((x, r))$ are open.

(a) Show that this determines a Hausdorff topology.

(b) Show that Γ is separable, but that R with the induced topology is not.

(c) Show that Γ is separable and first countable, but not second countable.

(d) Show that $Q = \{(x, 0) : x \in \mathbb{Q}\}$ and $I = \{(x, 0) : x \notin \mathbb{Q}\}$ are disjoint closed sets; but if $I \subset U$ and $Q \subset V$, where U and V are open, then $U \cap V \neq \varnothing$. ($\Gamma$ is not normal.)

(e) Show that if $C \subset \Gamma$ is closed and $p \notin C$, there is a continuous function $f : \Gamma \to [0, 1]$ so that $f|_C = 0$ and $f(p) = 1$. (Γ is completely regular.)

8. A filter on $\mathbb{N}$ is a non-empty collection $\mathcal{U}$ of non-empty subsets of $\mathbb{N}$ with the property that if $A, B \in \mathcal{U}$, then $A \cap B \in \mathcal{U}$ and if $C \supset A$, then $C \in \mathcal{U}$.

(a) Use Zorn's lemma to show that there exist a maximal filter $\mathcal{U}$ of $\mathbb{N}$ containing all co-finite sets. (This is known as a *free ultrafilter*.)

(b) Show that for each $A \subset \mathbb{N}$, either $A \in \mathcal{U}$ or $A^c \in \mathcal{U}$.

(c) Show that there is no subsequence of $\mathbb{N}$ such that $\lim_{i\to\infty} x_{n_i}$ exists for all bounded sequences $x = (x_n)$.

(d) For $A \in \mathcal{U}$, let $n_A = \min\{k : k \in A\}$. Prove that n_A is a (cofinal) subnet of the sequence $1, 2, 3, \dots$ with the property that $\lim_{A\in\mathcal{U}} x_{n_A}$ does exist for all bounded sequences $x = (x_n)$.

CHAPTER 2

Banach Spaces

2.1. Examples

2.1.1. DEFINITION. A *norm* on a vector space V over a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is a function $\|\cdot\| : V \to [0, \infty)$ such that

(1) $\|x\| = 0 \iff x = 0$ (definite)

(2) $\lambda x\| = |\lambda|\,\|x\|$ for all $\lambda \in \mathbb{F}$ and $x \in V$ (positive homogeneous)

(3) $\|x + y\| \le \|x\| + \|y\|$ (triangle inequality)

A *Banach space* is a complete normed vector space.

2.1.2. EXAMPLES.

(1) $C(X)$, the space of continuous complex valued functions on a compact Hausdorff space X with $\|f\|_\infty = \sup_{x\in X} |f(x)|$. A sequence is Cauchy in the sup norm precisely when it converges uniformly. The uniform limit of continuous functions is continuous, so $C(X)$ is complete. Variants include $C_{\mathbb{R}}(X)$, the real vector space of real valued continuous functions.

Also if X is a locally compact Hausdorff space, then $C_0(X)$ consists of continuous functions on X such that $K_\varepsilon = \{x : |f(x)| \ge \varepsilon\}$ is compact for all $\varepsilon > 0$. This is also a Banach space with the sup norm. Note that the space $C_c(X)$ of continuous functions on X with compact support, where. $\operatorname{supp}(f) = \overline{\{x : f(x) \ne 0\}}$, is a normed vector space with the sup norm, but it is not complete unless X is compact. The completion of $C_c(X)$ in the sup norm is $C_0(X)$.

A very important example in this family of spaces is

$$c_0 = \{(a_i)_{i\ge1} : \lim_{i\to\infty} a_i = 0\} \quad \text{with } \|(a_i)\|_\infty = \sup_{i\ge1} |a_i|.$$

(2) For $1 \le p < \infty$, $l_p = \{(a_i)_{i\ge1} : \|(a_i)\|_p = \big(\sum_{i\ge1} |a_i|^p\big)^{1/p} < \infty\}$ and $l_\infty = \{(a_i)_{i\ge1} : \|(a_i)\|_\infty = \sup |a_i| < \infty\}$. Likewise if $I \subset \mathbb{R}$ is an interval (finite or infinite), one can put the p-norm on $C_c(I)$ by $\|f\|_p = \Big(\int_I |f(x)|^p\,dx\Big)^{1/p}$. The completion is called $L^p(I)$ for $1 \le p < \infty$. Here the measure involved is Lebesgue measure on I.

In greater generality, if μ is a measure on a measure space $(X, \mathcal{B})$, then "$L^p(\mu)$" consists of all measurable functions on X such that $\|f\|_p = \big(\int_I |f|^p\,d\mu\big)^{1/p} < \infty$.

Then we define $L^p(\mu) = $ "$L^p(\mu)$"$/\mathcal{N}$ where $\mathcal{N} = \{f : f = 0\,\text{a.e.}(\mu)\}$. That is, elements of $L^p(\mu)$ are equivalence classes of functions. This is necessary so that the norm is definite. The space $L^\infty(\mu)$ consists of all bounded measurable functions modulo $\mathcal{N}$ with $\|f\|_\infty = \operatorname{ess\,sup}|f| = \sup\{r \geq 0 : \mu(|f| > r) > 0\}$. In all of these cases, the triangle inequality is called Minkowski's inequality. Completeness is a theorem of Riesz usually established in measure theory courses.

When $1 \leq p < \infty$, define q so that $\frac{1}{p} + \frac{1}{q} = 1$ ($q = \infty$ for $p = 1$). Hölder's inequality states that if $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then

$$\Big|\int fg\,d\mu\Big| \leq \|f\|_p\,\|g\|_q.$$

(3) $\qquad C^{(n)}[a, b] = \{f(x) : f \text{ has } n \text{ continuous derivatives on } [a, b]\}.$

Define $\|f\|_{C^{(n)}} = \sum_{k=0}^n \|f^{(k)}\|_\infty$. In this norm, a sequence $f_n \to f$ if and only if $f_n^{(k)}$ converges uniformly to $f^{(k)}$ for $0 \leq k \leq n$. The completeness of $C^{(n)}[a, b]$ follows from the fact that the uniform limit of continuous functions is continuous, and the fact that the Riemann integral of a uniform limit is the uniform limit of the integral, combined with the Fundamental Theorem of Calculus.

(4) Hilbert spaces. An inner product space is a vector space V over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ with a positive definite sesquilinear form (inner product). That is $\langle x, x\rangle \geq 0$ with equality only when $x = 0$; $\langle y, x\rangle = \overline{\langle x, y\rangle}$ and $\langle \lambda x, \mu y\rangle = \lambda\overline{\mu}\langle x, y\rangle$. A norm is defined by $\|x\| = \langle x, x\rangle^{1/2}$. A complete inner product space is called a *Hilbert space*. The Cauchy-Schwarz inequality is $|\langle x, y\rangle| \leq \|x\|\,\|y\|$. This is valid in all inner product spaces. The triangle inequality follows from

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

Examples of Hilbert spaces include l_2, $L^2(I)$ and $L^2(\mu)$ with $\langle f, g\rangle = \displaystyle\int f\bar{g}\,d\mu$.

A less obvious example is the following. Let Ω be an open subset of $\mathbb{C}$. Define $L^2_a(\Omega)$ to be the vector space of all analytic functions on Ω such that

$$\|f\|_2 = \Big(\int_\Omega |f(z)|^2\,dA\Big)^{1/2} < \infty$$

where $dA = dx\,dy$ is planar measure on Ω. Define $\langle f, g\rangle = \displaystyle\int_\Omega f(z)\overline{g(z)}\,dA$. The non-obvious fact is completeness.

Let $z_0 \in \Omega$. There is some $\rho > 0$ so that $\overline{b_\rho(z_0)} \subset \Omega$. Then using polar coordinates, ($dA = r\,dr\,d\theta$)

$$\int_{b_\rho(z_0)} f(z)\,dA = \int_0^\rho\int_0^{2\pi} f(z_0 + re^{i\theta})d\theta\,r\,dr = \int_0^\rho 2\pi f(z_0)d\theta\,r\,dr = \pi\rho^2 f(z_0).$$

Therefore by Cauchy-Schwarz,

$$\begin{aligned} |f(z_0)| &= \Big| \frac{1}{\pi\rho^2} \int_{b_\rho(z_0)} f(z)\, dA \Big| \\ &\le \frac{1}{\pi\rho^2} \Big(\int_{b_\rho(z_0)} |f(z)|^2\, dA \Big)^{1/2} \Big(\int_{b_\rho(z_0)} 1\, dA \Big)^{1/2} \\ &\le \frac{1}{\pi\rho^2} \|f\|_2 \sqrt{\pi\rho^2} = \frac{\|f\|_2}{\sqrt{\pi}\rho}. \end{aligned}$$

This says that evaluation of $f \in L^2_a(\Omega)$ at z_0 is continuous. (This is false for $L^2(I)$ where elements are only defined a.e.)

Now suppose that f_n is a Cauchy sequence in $L^2_a(\Omega)$. So for $\varepsilon > 0$, there is an N so that for $m, n \ge N$, $\|f_m - f_n\|_2 < \varepsilon$. Hence for

$$z \in \Omega_n = \{z \in \Omega : \text{dist}(z, \Omega^c) \ge \tfrac{1}{n}\},$$

we have $|f_n(z) - f_n(z)| \le \varepsilon/\sqrt{\pi}\rho$. This means that $f_n(z)$ is uniformly Cauchy on Ω_n, and thus converges uniformly there to a function $f(z)$. That is, $f_n \to f$ uniformly on compact subsets of Ω. Therefore the limit $f(z)$ is analytic. It is routine to check that $\|f\|_2 = \lim_{n\to\infty} \|f_n\|_2 < \infty$. So $f \in L^2_a(\Omega)$ and $f_n \to f$ in the $L^2_a(\Omega)$ norm.

2.2. Constructions of Banach Spaces

The following easy proposition is basic to functional analysis.

2.2.1. PROPOSITION. *Let X and Y be normed vector spaces, and let $T : X \to Y$ be a linear map. The following are equivalent:*

(1) $\|T\| := \sup_{\|x\| \le 1} \|Tx\| < \infty$. *i.e. T is* bounded.

(2) *T is Lipschitz, and hence uniformly continuous.*

(3) *T is continuous.*

(4) *T is continuous at $x = 0$.*

PROOF. (1)⇒(2). $\|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| \le \|T\|\,\|x_1 - x_2\|$. So F is Lipschitz with constant $\|T\|$. In particular, given $\varepsilon > 0$, if $\|x_1 - x_2\| < \varepsilon/\|T\|$, then $\|Tx_1 - Tx_2\| < \varepsilon$. So T is uniformly continuous.

(2)⇒(3)⇒(4) is trivial.

(4)⇒(1) is equivalent to ¬(1)⇒ ¬(4). So suppose that $\|T\| = \infty$. Then there are $x_n \in X$ with $\|x_n\| \le 1$ and $\|Tx_n\| > n^2$. Hence $\frac{1}{n}x_n \to 0$ but $\|T(\frac{1}{n}x_n)\| \to \infty$. So T is not continuous at $x = 0$. ∎

2.2.2. DEFINITION. Let X and Y be normed vector spaces. Then the space of *bounded linear operators* from X to Y is

$$\mathcal{B}(X,Y) = \{T : X \to Y \text{ continuous linear maps}\} \text{ with } \|T\| = \sup_{\|x\|\leq 1} \|Tx\|.$$

When $Y = X$, we write $\mathcal{B}(X)$ for $\mathcal{B}(X,X)$. When $Y = \mathbb{F}$ is the scalar field for X, we write X^* for $\mathcal{B}(X,\mathbb{F})$. This is called the *dual space* of X. Elements of X^* are called (continuous linear) *functionals*.

It is easy to check that $\|\cdot\|$ is a norm. This is left to the reader.

2.2.3. PROPOSITION. *Let X be a normed vector space and let Y be a Banach space. Then $\mathcal{B}(X,Y)$ is a Banach space.*

PROOF. We will verify completeness. Let $(T_n)_{n\geq 1}$ be a Cauchy sequence in $\mathcal{B}(X,Y)$. For $\varepsilon > 0$, there is an N_ε so that if $N_\varepsilon \leq n < m$, then $\|T_m - T_n\| < \varepsilon$. Hence for each $x \in X$,

$$\|T_m x - T_n x\| \leq \|T_m - T_n\|\,\|x\| < \varepsilon\|x\|.$$

This means that $(T_n x)$ is a Cauchy sequence in Y for every $x \in X$. Since Y is complete, we may define $Tx := \lim_{n\to\infty} T_n x$. It is easy to check that T is linear. Moreover if $n \geq N_\varepsilon$ and $\|x\| \leq 1$,

$$\|Tx - T_n x\| = \lim_{m\to\infty} \|T_m x - T_n x\| \leq \varepsilon.$$

It follows that $\|T\| \leq \|T_{N_\varepsilon}\| + \varepsilon < \infty$ and that $\|T - T_n\| \to 0$ and $n \to \infty$. Thus $\mathcal{B}(X,Y)$ is complete. ∎

2.2.4. THEOREM. *Let $1 < p < \infty$ and let q satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Then the map taking $y = (y_j) \in l_q$ to the linear functional $\varphi_y(x) = \sum_{j\geq 1} x_j y_j$ is an isometric map onto l_p^*.*

PROOF. Let e_j denote the element of l_p with a 1 in the jth entry and 0s elsewhere. If $y = (y_j) \in l_q$ and $x = \sum_{j=1}^N x_j e_j$, then $\varphi_y(x) := \sum_{j=1}^N x_j y_j$. By Hölder's inequality,

$$|\varphi_y(x)| \leq \sum_{j=1}^N |x_j|\,|y_j| \leq \|x\|_p\,\|y\|_q.$$

This shows that $\|\varphi_y\| \leq \|y\|_q$ on the span$\{e_j : j \geq 1\}$, and so it extends by continuity to all of l_p, and indeed, $\varphi_y(x) := \sum_{j=1}^\infty x_j y_j$ converges absolutely for all $x \in l_p$. This shows that $\|\varphi_y\| \le \|y\|_q$. Taking $x_j = \overline{\text{sign}(y_j)}|y_j|^{q/p}$, we obtain

$$\|x\|_p^p = \sum_{j\ge 1} |y_j|^q = \|y\|_q^q < \infty;$$

so that $x \in l_p$ with $\|x\|_p = \|y\|_q^{q/p}$. Evaluating φ_y at $x/\|x\|_p$ shows

$$\begin{aligned}\|\varphi_y\| \ge \frac{|\varphi_y(x)|}{\|x\|_p} &= \frac{1}{\|y\|_q^{q/p}} \sum_{j\ge 1} |y_j|^{q/p}\, |y_j|^{q/q} \\ &= \frac{1}{\|y\|_q^{q/p}} \sum_{j\ge 1} |y_j|^q = \|y\|_q^{q(1-1/p)} = \|y\|_q.\end{aligned}$$

Therefore the map $y \to \varphi_y$ is an isometric map of l_q into l_p^*.

If $\varphi \in l_p^*$, let $y_j = \varphi(e_j)$. Again we define $x_j = \overline{\text{sign}(y_j)}|y_j|^{q/p}$. For any finite N, we have $u_N = \sum_{j=1}^N x_j e_j \in l_p$ and $v_N = \sum_{j=1}^N y_j f_j \in l_q$, where f_j is the sequence in l_q with a 1 in the jth entry and 0s elsewhere. By the previous paragraph, $\|u_N\|_p = \|v_N\|_q^{q/p}$ and

$$\|\varphi\| \ge \frac{|\varphi(u_N)|}{\|u_N\|_p} = \|v_N\|_q.$$

Hence $y = (y_j)$ satisfies

$$\|y\|_q = \lim_{N\to\infty} \|v_N\|_q \le \|\varphi\| < \infty.$$

So $y \in l_q$. Moreover for any $x = \sum_{j=1}^N x_j e_j \in \text{span}\{e_j : j \ge 1\}$, we have

$$\varphi(x) = \sum_{j=1}^N x_j \varphi(e_j) = \sum_{j=1}^N x_j y_j = \varphi_y(x).$$

Thus the two continuous functionals φ and φ_y agree on a dense subset of l_p, and thus they are equal. This shows that the isometry from l_q into l_p^* is surjective. ■

2.2.5. REMARKS. Theorem 2.2.4 is a very special case of a theorem of Riesz which is established in a course on measure theory. Namely that if $1 < p < \infty$, then $L^p(\mu)^* = L^q(\mu)$. This is also true for $p = 1$ (with $q = \infty$) provided that μ is σ-finite.

The other big theorem on dual spaces from measure theory is the Riesz Representation Theorem. If X is a locally compact Hausdorff space, then $C_0(X)^* = M(X)$, the space of complex (finite) regular Borel measures on X.

2.2.6. EXAMPLE. $c_0^* = l_1$. Again let e_j denote the sequence with a 1 in the jth coordinate and 0s elsewhere; and let f_j be the corresponding basis in l_1. If $\varphi \in c_0^*$, let $a_j = \varphi(e_j)$. Then

$$\|\varphi\| \geq \varphi\big(\tfrac{\overline{a_1}}{|a_1|}, \dots, \tfrac{\overline{a_j}}{|a_j|}, \dots, \tfrac{\overline{a_n}}{|a_n|}, 0, 0 \dots, \big) = \sum_{j=1}^{n} |a_j|.$$

Letting $n \to \infty$, we deduce that $y = (a_j)$ belongs to l_1 and $\|y\|_1 \leq \|\varphi\|$. Moreover $\varphi(x) = \varphi_y(x) = \sum_{j \geq 1} x_j a_j$. Conversely, each $y \in l_1$ determines a linear functional with $\|\varphi_y\| = \|y\|_1$.

Now consider $\mathcal{B}(c_0)$. If $T \in \mathcal{B}(c_0)$, define an infinite matrix $\big[t_{ij}\big]$ by $Te_j = \sum_{i \geq 1} t_{ij} e_i$. The columns Te_j belong to c_0, so $\lim_{i \to \infty} t_{ij} = 0$. Also the row $(t_{i1}, t_{i2}, t_{i3}, \dots)$ acts on a vector $x \in c_0$ by $\varphi_i(x) = \sum_{j \geq 1} t_{ij} x_j = \varphi_{f_i}(Tx)$. This determines the ith coordinate of $Tx = \sum_{i \geq 1} \varphi_i(x) e_i$. So each $\varphi_i = \varphi_{f_i} \circ T$ is a continuous linear functional on c_0, so there is a vector $y_i \in l_1$ so that $\varphi_i = \varphi_{y_i}$. Moreover $\|\varphi_i\| = \|y_i\|_1 \leq \|T\|$. In fact,

$$\mathcal{B}(c_0) = \big\{T = \big[t_{ij}\big] : (t_{ij})_{i \geq 1} \in c_0 \text{ for } j \geq 1 \text{ and } \sup_{i \geq 1} \sum_{j \geq 1} |t_{ij}| < \infty\big\}.$$

To see this, note that the boundedness of the rows in the l_1 norm ensures that T is a bounded linear map from c_0 into l_∞. The condition on the columns ensures that $Te_j \in c_0$ for $g \geq 1$. So $T \operatorname{span}\{e_j : j \geq 1\} \subset c_0$. The subspace $\operatorname{span}\{e_j : j \geq 1\}$ is dense in c_0 and T is continuous. Thus, since c_0 is closed in l_∞, the range of T is contained in c_0. Therefore $T \in \mathcal{B}(c_0)$.

This is a rather special property of c_0. There are very few Banach spaces where the space of bounded linear maps can be simply characterized.

An easy way to obtain a new Banach space from another is to take a *closed subspace*, which as the term suggests is a subspace which is norm closed. Since a closed subset of a complete metric space is complete, a closed subspace of a Banach space is a Banach space.

2.2.7. Example. Let $A(\mathbb{D}) = \{f \in C(\mathbb{T}) : \hat{f}(n) = 0 \text{ for } n < 0\}$. Here $\mathbb{T}$ is the unit circle $\mathbb{T} = \{e^{i\theta} : \theta \in [-\pi, \pi]\}$. The (complex) Fourier coefficients of $f \in C(\mathbb{T})$ are given by

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta}\, d\theta \quad \text{for} \quad n \in \mathbb{Z}.$$

The linear functionals $\varphi_n(f) = \hat{f}(n)$ are continuous since $|\hat{f}(n)| \leq \|f\|_\infty$, so $\|\varphi_n\| \leq 1$. They are exactly norm 1 since $\|e^{in\theta}\|_\infty = 1$ and $\varphi_n(e^{in\theta}) = 1$. Hence the kernel, $\ker \varphi_n$, is a closed subspace. It follows that $A(\mathbb{D}) = \bigcap_{n<0} \ker \varphi_n$ is a closed subspace of $C(\mathbb{T})$.

Every continuous function $f \in C(\mathbb{T})$ extends uniquely to a harmonic function on the open disk by

$$\begin{aligned}\tilde{f}(re^{i\theta}) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})P_r(\theta - t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})\frac{1-r^2}{1-2r\cos(\theta-t)+r^2}\,dt \\ &= \sum_{n=-\infty}^{\infty} \hat{f}(n)r^{|n|}e^{in\theta}.\end{aligned}$$

Poisson's Theorem shows that $\tilde{f}(re^{i\theta})$ converges uniformly to $f(e^{i\theta})$ as $r \to 1$. Hence $\tilde{f}$ extends to be continuous on the closed disk, and coincides with f on the boundary circle. For $f \in A(\mathbb{D})$, all negative Fourier coefficients vanish. Thus, writing $z = re^{i\theta}$, we get that $\tilde{f}(z) = \sum_{n\geq 0}\hat{f}(n)z^n$. This power series converges on the open unit disk $\mathbb{D}$, and converges uniformly when $|z| \leq r < 1$. So for $f \in A(\mathbb{D})$, $\tilde{f}(z)$ is analytic on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$. Conversely every function which has these properties has a power series which converges on the open disk. By Cauchy's Theorem, $f(z)$ for $z \in \mathbb{D}$ is determined by the boundary values. Indeed, analytic functions are harmonic, so this is the Poisson extension of the boundary values. It is clear that each $f_r(e^{i\theta})$ lies in $A(\mathbb{D})$ since the Fourier series converges uniformly. They converge uniformly to $f(e^{i\theta})$ by continuity on $\overline{\mathbb{D}}$, and so $f \in A(\mathbb{D})$.

The third construction in this section is the notion of a quotient space.

2.2.8. DEFINITION. Let X be a Banach space and let M be a closed subspace. Then X/M is a vector space. Define a norm on X/M by

$$\|x + M\| = \inf\{\|x - m\| : m \in M\}.$$

2.2.9. PROPOSITION. *Let X be a Banach space and let M be a closed subspace. Then X/M is a Banach space.*

PROOF. First check that this is a bona fide norm. Note that $\|x + M\| = 0$ only if there are $m_n \in M$ with $\|x - m_n\| \to 0$. That means that x lies in $\overline{M} = M$. So $x + M = \dot{0}$ is the zero element. Positive homogeneity and the triangle inequality are straightforward.

For completeness, let $\dot{x}_n = x_n + M$ for $n \geq 1$ be a Cauchy sequence in X/M. Find a subsequence $\dot{x}_{n_i}$ so that $\|\dot{x}_{n_i} - \dot{x}_{n_{i+1}}\| < 2^{-i}$. Recursively choose $m_i \in M$ so that $x'_{n_i} = x_{n_i} - m_i$ satisfy $\|x'_{n_i} - x'_{n_{i+1}}\| < 2^{-i}$. Then $(x'_{n_i})_{i\geq 1}$ is Cauchy in X. Let x be the limit. Then it is clear that $\dot{x}_{n_i}$ converge to $\dot{x}$. A standard argument now shows that the whole sequence converges to $\dot{x}$. Hence X/M is complete, and thus is a Banach space. ■

2.2.10. PROPOSITION. *Let X be a Banach space and let M be a proper closed subspace. Define a linear map $Q : X \to X/M$ by $Qx = \dot{x}$. Then $Q \in \mathcal{B}(X, X/M)$ is surjective, $\|Q\| = 1$ and* $\ker Q = M$.

PROOF. It is clear that Q is a linear surjection with $\ker Q = M$. Moreover, since $\|\dot{x}\| \le \|x\|$, we have $\|Q\| \le 1$. Since M is proper, pick $x \in X \setminus M$; so that $\dot{x} \ne \dot{0}$. Replace x by $x/\|\dot{x}\|$ so that $1 = \|\dot{x}\| = \inf\{\|x - m\| : m \in M\}$. Choose $m_n \in M$ so that $\|x - m_n\| \to 1$. Then $\|Q\| \ge \sup_{n \ge 1} \frac{\|\dot{x}\|}{\|x - m_n\|} = 1$. Therefore $\|Q\| = 1$. ∎

The map $Q : X \to X/M$ is called the *quotient map*.

2.2.11. EXAMPLE. Let X be a compact Hausdorff space, and let E be a closed subset of X. Then $I(E) = \{f \in C(X) : f|_E = 0\}$ is a closed subspace (and in fact, an ideal). Consider $C(X)/I(E)$.

Let $h \in C(X)$. Then

$$\|\dot{h}\| = \inf\{\|h - f\|_\infty : f|_E = 0\} \ge \sup_{x \in E} |h(x)| = \|h|_E\|_\infty.$$

On the other hand, suppose that $\|h|_E\|_\infty = 1$. Define a continuous function $g : \mathbb{C} \to \overline{\mathbb{D}}$ by $g(z) = \begin{cases} z & \text{if } |z| \le 1 \\ \frac{z}{|z|} & \text{if } |z| > 1 \end{cases}$. Then $g(h(x))$ has norm 1, and $g|_E = h|_E$. Then $f = h - g \circ h$ vanishes on E, so $f \in I(E)$. Moreover $h - f = g$, so that $\|\dot{h}\| \le \|g\|_\infty = 1$. We deduce that $\|\dot{h}\| = \|h|_E\|_\infty$.

Let $Q : C(X) \to C(X)/I(E)$ be the quotient map. Define a linear map $T : C(X)/I(E) \to C(E)$ by $T\dot{h} = h|_E$. This makes sense since $h - f|_E = h|_E$ for all $f \in I(E)$. Moreover we have just shown that $\|T\dot{h}\| = \|\dot{h}\|$ for all $h \in C(X)/I(E)$. That is, T is an *isometry*. Observe that $TQ = R$, where R is the restriction map $Rh = h|_E$ from $C(X)$ onto $C(E)$. Also T maps onto $C(E)$ by Tietze's Extension Theorem. So $C(X)/C(E)$ is *isometrically isomorphic* tp $C(E)$.

The final construction in this section is the notion of a direct sum.

2.2.12. DEFINITION. Let X and Y be Banach spaces. Consider the *direct sum* $X \oplus_p Y = \{(x, y) : x \in X, y \in Y\}$. For $1 \le p \le \infty$, let $X \oplus_p Y$ endow the direct sum with the norm $\|(x, y)\|_p = \left(\|x\|^p + \|y\|^p\right)^{1/p}$ if $p < \infty$ and $\|(x, y)\|_\infty = \max\{\|x\|, \|y\|\}$.

The following result is straightforward, and the proof is left to the reader.

2.2.13. PROPOSITION. *Let X and Y be Banach spaces. Then $X \oplus_p Y$ is a Banach space for $1 \le p \le \infty$. Moreover*

$$\|(x,y)\|_\infty \le \|(x,y)\|_p \le \|(x,y)\|_1 \le 2\|(x,y)\|_\infty.$$

We say that two norms on X are *equivalent* if there are constanst $0 < C_1 \le C_2 < \infty$ so that $C_1\|x\|_2 \le \|x\|_1 \le C_2\|x\|_2$ for all $x \in X$.

2.2.14. PROPOSITION. *Let X be a finite dimensional vector space. Then any two norms on X are equivalent. In particular, X is complete in any norm. So every finite dimensional subspace of a Banach space is closed.*

PROOF. Let $\mathbb{F}$ be the field of scalars for X. Fix a basis $x_1, \dots, x_n$ for X, and consider a norm $\|\cdot\|$ on X. Define a linear isomorphism $T : (\mathbb{F}^n, \|\cdot\|_2) \to X$ by $T\sum_{i=1}^n a_i e_i = \sum_{i=1}^n a_i x_i$. Then for $u = \sum_{i=1}^n a_i e_i \in \mathbb{F}^n$,

$$\|Tu\| = \Big\| \sum_{i=1}^n a_i x_i \Big\| \le \sum_{i=1}^n |a_i|\, \|x_i\| \le \|u\|_2 \Big(\sum_{i=1}^n \|x_i\|^2 \Big)^{1/2} = C\|u\|_2$$

by the Cauchy-Schwarz inequality. Therefore T is a continuous linear bijection.

Let $S = \{u \in \mathbb{F}^n : \|u\|_2 = 1\}$ be the unit sphere. This is compact, and therefore TS is compact in X. Since $0 \notin TS$, the continuous function $f(x) = \|x\|$ attains a minimum value on TS, say $\min_{u\in S} \|Tu\| = r > 0$. It follows that $T^{-1}b_r(X) \subset b_1(\mathbb{F}^n)$; for if $\|T^{-1}x\| = t \ge 1$, then $x/t \in b_r(X)$ and $T^{-1}(x/t)$ would belong to S So $\|T^{-1}\| \le \frac{1}{r}$. Therefore T is an isomorphism, and hence a biLipschitz homeomorphism. It follows that $(X, \|\cdot\|)$ is complete, and this norm is equivalent to the Euclidean norm $\|x\|_2 := \|T^{-1}x\|_2$.

If $|\!|\!|\cdot|\!|\!|$ is another norm on X, then it is also equivalent to $\|\cdot\|_2$, and thus is equivalent to $\|\cdot\|$. If X is a finite dimensional subspace of a Banach space Y, then X is complete in this norm, and hence it is closed in Y. ■

2.2.15. PROPOSITION. *The closed unit ball of a Banach space is compact if and only if it is finite dimensional.*

PROOF. If X has finite dimension n, then the norm is equivalent to the Euclidean norm on $\mathbb{F}^n$. Thus every closed bounded set is compact.

If X has infinite dimension, we can recursively choose a sequence of unit vectors $(x_n)_{n\ge1}$ so that $\operatorname{dist}(x_{n+1}, \operatorname{span}\{x_1, \dots, x_n\}) > \frac{1}{2}$. Suppose that $x_1, \dots, x_n$ are chosen, and let $M_n = \operatorname{span}\{x_1, \dots, x_n\}$. Then M_n is closed, and X/M_n is an infinite dimensional Banach space. Let $\dot{x} = x + M_n \in X/M_n$ with

$$\tfrac{1}{2} < \|\dot{x}\| = \inf_{m\in M_n} \|x+m\| < 1.$$

Choose $x+m$ so that $\|x+m\| < 1$; and define $x_{n+1} = \frac{x+m}{\|x+m\|}$. Then there is no finite $\frac{1}{4}$-net in the unit ball since a point can be within $\frac{1}{4}$ of only a single x_n. Therefore $\overline{b_1(X)}$ is not compact. ∎

2.3. Hilbert spaces

2.3.1. DEFINITION. In an inner product space, $x \perp y$ means $\langle x, y\rangle = 0$ and the vectors are *orthogonal*. A set $\{e_\alpha : \alpha \in A\}$ is *orthonormal* if $\langle e_\alpha, e_\beta\rangle = \delta_{\alpha\beta}$. If $S \subset H$, then $S^\perp = \{x \in H : x \perp y \text{ for all } y \in S\}$.

If $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$, the Pythagorean law. Therefore if $e_1, \dots, e_n$ are orthonormal, then $\big\| \sum_{i=1}^n a_i e_i \big\|^2 = \sum_{i=1}^n |a_i|^2$.

2.3.2. DEFINITION. A series $\sum_{\alpha\in A} a_\alpha e_\alpha$ *converges unconditionally* in a Hilbert space if for the upward directed poset Λ of finite subsets $F \subset A$ (ordered by inclusion), the net $x_F = \sum_{\alpha\in F} a_\alpha e_\alpha$ converges.

2.3.3. THEOREM. *Let H be a Hilbert space, let $\{e_\alpha : \alpha \in A\}$ be a non-empty orthonormal set in H, and let $M = \overline{\text{span}\{e_\alpha : \alpha \in A\}}$. Then*

(1) *The set $\{e_\alpha : \alpha \in A\}$ is linearly independent; and*

$$\text{dist}\big(e_\alpha, \overline{\text{span}\{e_\beta : \beta \neq \alpha\}}\big) = 1.$$

(2) *(Bessel's inequality) If $x \in H$ and $a_\alpha = \langle x, e_\alpha\rangle$, then $\sum\limits_{\alpha\in A} |a_\alpha|^2 \le \|x\|^2$.*

(3) *If $a_\alpha \in \mathbb{C}$ and $\sum_{\alpha\in A} |a_\alpha|^2 = L^2 < \infty$, then $\sum_{\alpha\in A} a_\alpha e_\alpha$ converges unconditionally to a vector x with $\|x\| = L$. Moreover $\langle x, e_\alpha\rangle = a_\alpha$.*

(4) *If $x \in H$, let $a_\alpha = \langle x, e_\alpha\rangle$ and define $Px = \sum_{\alpha\in A} a_\alpha e_\alpha$. Then P is in $\mathcal{B}(H)$ with $||P|| = 1$, $P^2 = P$ and $P(H) = M$.*

(5) $\ker P = M^\perp$. *Also $Px = x$ if and only if $x \in M$. And $M + M^\perp = H$.*

(6) *(Parsevel's identity): If $y \in M$, then $\|y\| = \big(\sum_{\alpha\in A} |\langle y, e_\alpha\rangle|^2\big)^{1/2}$.*

The operator P constructed here is the *orthogonal projection* of H onto M.

PROOF. (1) Elements of $\text{span}\{e_\alpha : \alpha \in A\}$ (before closure) have the form $\sum_{\beta\in F} a_\beta e_\beta$ for $F \subset A$ finite. By the Pythagorean law, if $F \subset A \setminus \{\alpha\}$ is finite,

$$\|e_\alpha - \sum_{\beta\in F} a_\beta e_\beta\|^2 = 1 + \sum_{\beta\in F} |a_\beta|^2.$$

Thus the minimum value, 1, is obtained when $a_\beta = 0$ for all $\beta \neq \alpha$. This extends to the closure by continuity of the norm. In particular, $\{e_\alpha : \alpha \in A\}$ are linearly independent.

(2) For $F \subset A$ finite, let $x_F = \sum_{\alpha \in F} a_\alpha e_\alpha$. Then

$$\begin{aligned} 0 \le \|x - x_F\|^2 &= \|x\|^2 - 2\,\mathrm{Re}\langle x, x_F\rangle + \|x_F\|^2 \\ &= \|x\|^2 - 2\sum_{\alpha \in F} |a_\alpha|^2 + \sum_{\alpha \in F} |a_\alpha|^2 = \|x\|^2 - \sum_{\alpha \in F} |a_\alpha|^2. \end{aligned}$$

Taking the sup over $F \subset A$ yields $\sum_{\alpha \in A} |a_\alpha|^2 \le \|x\|^2$.

(3) Again define $x_F = \sum_{\alpha \in F} a_\alpha e_\alpha$. Then $\|x_F\|^2 = \sum_{\alpha \in F} |a_\alpha|^2 \le L^2$. Note that $L^2 = \sup_{\substack{F \subset A,\\ F \text{finite}}} \|x_F\|^2$. Given $\varepsilon > 0$, choose $F_0 \subset A$ finite so that $\|x_{F_0}\|^2 > L^2 - \varepsilon$. Then for $F, G \ge F_0$,

$$\|x_F - x_G\|^2 = \sum_{\alpha \in F \triangle G} |a_\alpha|^2 \le \sum_{\alpha \in A \setminus F_0} |a_\alpha|^2 < L^2 - (L^2 - \varepsilon) = \varepsilon.$$

Therefore $\{x_F\}_\Lambda$ is a Cauchy net. Since H is complete, it has a limit $x \in H$. Since each $x_F \in M$ and M is closed, $x \in M$. Hence $\|x\|^2 = \lim_\Lambda \|x_F\|^2 = L^2$. Compute $\langle x, e_\alpha\rangle = \lim_\Lambda \langle x_F, e_\alpha\rangle = a_\alpha$ since $\langle x_F, e_\alpha\rangle = a_\alpha$ when $F \ge \{\alpha\}$.

(4) By (2) and (3), for each $x \in H$ with $a_\alpha = \langle x, e_\alpha\rangle$, the sum $Px = \sum_{\alpha \in A} a_\alpha e_\alpha$ is well defined and lies in M. The map P is evidently linear. Moreover $\|Px\|^2 = \sum_{\alpha \in A} |a_\alpha|^2 \le \|x\|^2$ by (3) and Bessel's inequality. Therefore $\|P\| \le 1$. One can check that $Pe_\alpha = e_\alpha$, and thus $\|P\| = 1$. (3) also shows that $\langle Px, e_\alpha\rangle = a_\alpha$. Therefore $P^2 x = Px$. The range of P contains the unclosed span$\{e_\alpha : \alpha \in A\}$, and $P \sum_{\alpha \in F} a_\alpha e_\alpha = \sum_{\alpha \in F} a_\alpha e_\alpha$. That is, P is the identity on this subspace. By continuity, $Py = y$ for all $y \in M$.

(5) $Px = 0$ iff $\langle x, e_\alpha\rangle = 0$ for all $\alpha \in A$ iff $\langle x, \sum_{\alpha \in F} a_\alpha e_\alpha\rangle = 0$ for all vectors in span$\{e_\alpha : \alpha \in A\}$ iff $\langle x, y\rangle = 0$ for all $y \in M$ by continuity iff $x \in M^\perp$. We showed that $Py = y$ for $y \in M$, and since $Px \in M$, we have $Px = x$ only if $x \in M$. If $x \in H$, then $x = Px + (I - P)x$, and $Px \in M$ and $P(I - P)x = (P - P^2)x = 0$, so $(I - P)x \in M^\perp$. Hence $M + M^\perp = H$.

(6) This is immediate from (3) and (5). ∎

2.3.4. DEFINITION. An *orthonormal basis* for a Hilbert space H is an orthonormal set $\{e_\alpha : \alpha \in A\}$ such that $\overline{\mathrm{span}\{e_\alpha : \alpha \in A\}} = H$.

2.3.5. PROPOSITION. *Every Hilbert space has an orthonormal basis.*

PROOF. Order the collectiion $\mathcal{O}$ of all orthonormal sets by inclusion. Then $\mathcal{O}$ is inductive because if $\mathcal{C}$ is a chain of orthonormal sets, then the union is an orthonormal set containing every element of $\mathcal{C}$. By Zorn's Lemma, there is a maximal orthonormal set, say $\{e_\alpha : \alpha \in A\}$. Let $M = \overline{\mathrm{span}\{e_\alpha : \alpha \in A\}}$. If $M = H$, we are done. Otherwise $M^\perp \neq \{0\}$. Let f be a unit vector in $M^\perp$. Then $\{f, e_\alpha : \alpha \in A\}$ is a larger orthonormal set, contradicting the maximality of $\{e_\alpha : \alpha \in A\}$. Hence we must have $M = H$. ■

2.3.6. THEOREM. *Let H be a Hilbert space. For each $\varphi \in H^*$, there is a unique vector $y \in H$ so that $\varphi(x) = \langle x, y\rangle$. The map $\Upsilon : H^* \to H$ given by $\Upsilon(\varphi) = y$ is a conjugate linear isometry of H^* onto H. In particular, H^* is a Hilbert space with the inner product $\langle \varphi, \psi\rangle = \overline{\langle \Upsilon\varphi, \Upsilon\psi\rangle}$. The set $\{\Upsilon^{-1} e_\alpha : \alpha \in A\}$ forms an orthonormal basis for H^*.*

PROOF. Let $\{e_\alpha : \alpha \in A\}$ be an orthonormal basis for H. Define $a_\alpha = \varphi(e_\alpha)$. For every finite set $F \subset A$, let $y_F = \sum_{\alpha\in F} \overline{a_\alpha} e_\alpha$. Then

$$|\varphi(y_F)| = \sum_{\alpha\in F} |a_\alpha|^2 \le \|\varphi\|\,\|y_F\| = \|\varphi\| \Big(\sum_{\alpha\in F} |a_\alpha|^2 \Big)^{1/2}.$$

Hence $\sum_{\alpha\in F} |a_\alpha|^2 \le \|\varphi\|^2$. Taking the supremum over all finite subsets F yields $\sum_{\alpha\in A} |a_\alpha|^2 \le \|\varphi\|^2$. Therefore $y = \sum_{\alpha\in A} \overline{a_\alpha} e_\alpha$ belongs to H by Theorem 2.3.3, and $\|y\| \le \|\varphi\|$.

If $x = \sum_{\alpha\in F} b_\alpha e_\alpha$, then

$$\varphi(x) = \sum_{\alpha\in F} b_\alpha a_\alpha = \langle x, y\rangle.$$

Therefore φ agrees with the linear functional $\psi(x) = \langle x, y\rangle$ on the algebraic span of the basis, which is dense in H. Both are continuous, so they are equal. Note that $|\varphi(x)| = |\langle x, y\rangle| \le \|x\|\,\|y\|$. Take the supremum over all $x \in H$ with $\|x\| \le 1$ to get $\|\varphi\| \le \|y\|$. Hence $\|y\| = \|\varphi\|$. The uniqueness of y follows from the fact that any vector determining the functional φ must satisfy

$$\langle y, e_\alpha\rangle = \overline{\langle e_\alpha, y\rangle} = \overline{\varphi(e_\alpha)} = \overline{a_\alpha}.$$

By Theorem 2.3.3, $y = \sum_{\alpha\in A} \overline{a_\alpha} e_\alpha$.

Therefore there is a map $\Upsilon : H^* \to H$ given by $\Upsilon\varphi = y$. It is isometric because $\|\varphi\| = \|y\|$. It is surjective because for each $y \in H$, $\psi(x) = \langle x.y\rangle$ belongs to H^* and $\Upsilon\varphi = y$. However this is not a linear map. If $\psi \in H^*$ with $\Upsilon\psi = z$ and $\lambda, \mu \in \mathbb{C}$, then

$$(\lambda\varphi + \mu\psi)(x) = \lambda\langle x, y\rangle + \mu x, z = \langle x, \bar\lambda y + \bar\mu z\rangle.$$

Thus

$$\Upsilon(\lambda\varphi + \mu\psi) = \bar\lambda y + \bar\mu z = \bar\lambda \Upsilon\varphi + \bar\mu \Upsilon\psi.$$

This shows that Υ is a conjugate linear map.

Now define an inner product on H^* by $\langle \varphi, \psi\rangle = \overline{\langle \Upsilon\varphi, \Upsilon\psi\rangle}$. It is easy to check that this satisfies the properties of an inner product including linearity in the first variable. It is also clear that $\{\Upsilon^{-1}e_\alpha : \alpha \in A\}$ is orthonormal. Moreover since Υ is isometric,

$$\overline{\operatorname{span}\{\Upsilon^{-1}e_\alpha : \alpha \in A\}} = \Upsilon^{-1}\overline{\operatorname{span}\{e_\alpha : \alpha \in A\}} = \Upsilon^{-1}H = H^*.$$

So $\{\Upsilon^{-1}e_\alpha : \alpha \in A\}$ is an orthonormal basis. ■

2.3.7. DEFINITION. The *dimension* of a Hilbert space H is the cardinality of an orthonormal basis.

2.3.8. PROPOSITION. *The dimension of a Hilbert space is well defined.*

PROOF. We need to show that any two orthonormal bases have the same cardinality. If H is finite dimensional, this is just the cardinality of a basis, which is shown to be well defined in linear algebra.

If H is infinite dimensional, let $\{e_\alpha : \alpha \in A\}$ and $\{f_\beta : \beta \in B\}$ be two orthonormal cases. For each $\alpha \in A$, let $B_\alpha = \{\beta \in B : \langle e_\alpha, f_\beta\rangle \neq 0\}$. Then because

$$1 = \|e_\alpha\|^2 = \sum_{\beta \in B_\alpha} |\langle e_\alpha, f_\beta\rangle|^2,$$

it follows that B_α is non-empty and countable or finite. Indeed, there can be at most n βs such that $|\langle e_\alpha, f_\beta\rangle|^2 \geq \frac{1}{n}$. Thus there are at most countably many in B_α. Also $B = \bigcup_{\alpha\in A} B_\alpha$ because similarly every f_β has non-zero inner product with some e_α. Therefore $|B| \leq |A|\aleph_0 = |A|$. Similarly $|A| \leq |B|$. So by the Schroeder-Bernstein Theorem, $|A| = |B|$. ■

The following easy result is left as an exercise.

2.3.9. PROPOSITION. *Two Hilbert spaces are (isometrically) isomorphic if and only if they have the same dimension.*

2.3.10. REMARK. If H and K are Hilbert spaces, then $H \oplus_2 K$ is also a Hilbert space. We will write this as $H \oplus K$. Then H and K are orthogonal complements in the direct sum.

2.3.11. DEFINITION. A *unitary operator* $U \in \mathcal{B}(H, K)$ from one Hilbert space H to another, K, is a surjective isometry.

Note that U is a linear isomorphism which preserves norm. That means that is also preserves the inner product. This follows from the *polarization identity*:

$$\langle x, y\rangle = \frac{1}{4}\sum_{\varepsilon\in\{\pm1,\pm i\}} \varepsilon\|x + \varepsilon y\|^2 \quad \text{or} \quad \langle x, y\rangle = \frac{1}{2}\big(\|x + y\|^2 - \|x - y\|^2\big)$$

in the complex and real cases, respectively. Since U preserves norms, it preserves the inner product as well.

2.3.12. EXAMPLE. Let $H = L^2(\mathbb{T})$, the space of square integrable functions (modulo almost everywhere equality) with respect to Lebesgue measure $\frac{1}{2\pi}\,d\theta$. For $n \in \mathbb{Z}$, let $e_n(z) = z^n = e^{in\theta}$. Note that

$$\langle e_n, e_m\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{in\theta}\overline{e^{im\theta}}\,d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)\theta}\,d\theta = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}.$$

So $\{e_n : n \in \mathbb{Z}\}$ is an orthonormal set.

It is a fact that the trigonometric polynomials of the form $\sum\limits_{n=-N}^{N} a_n e^{in\theta}$ for $a_n \in \mathbb{C}$ and $N \in \mathbb{N}$ is uniformly dense in $C(\mathbb{T})$. This is a classical theorem of Weierstrass, which also follows from Féjer's Theorem or the Stone-Weierstrass Theorem. It is also a fact from measure theory that $C(\mathbb{T})$ is dense in $L^2(\mathbb{T})$ in the 2-norm. This follows from Lusin's Theorem. Hence $\{e_n : n \in \mathbb{Z}\}$ has dense linear span. Therefore $\{e_n : n \in \mathbb{Z}\}$ is an orthonormal basis.

Consider the linear map $U : L^2(\mathbb{T}) \to l_2(\mathbb{Z})$ given by $Uf = \big(\hat{f}(n)\big)_{n\geq 1}$ that takes a function to it Fourier series:

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{-in\theta}\,d\theta = \langle f, e_n\rangle.$$

This is the Fourier operator. This map takes the orthonormal basis $\{e_n : n \in \mathbb{Z}\}$ onto the standard basis $\{\delta_n : n \in \mathbb{Z}\}$ for $l_2(\mathbb{Z})$. The Parseval identity shows that U is an isometry. The range is therefore closed, and is dense, so is surjective. Thus U is unitary.

2.4. Category Theorems

2.4.1. DEFINITION. A subset $A \subset X$ of a topological space X is *nowhere dense* if $\overline{A}$ has no interior. A subset of a complete metric space is *first category* if it is the countable union of nowhere dense sets.

Recall the following important result from real analysis.

2.4.2. BAIRE CATEGORY THEOREM. *A non-empty complete metric space X is not a countable union of nowhere dense sets. Indeed, if U_n are dense open subsets of a complete metric space X, then $\bigcap_{n\geq 1} U_n$ is dense in X.*

PROOF. Note that $U_n = \overline{A_n}^c$ is a dense open set, and $\big(\bigcup_{n\geq 1} \overline{A_n}\big)^c = \bigcap_{n\geq 1} U_n$. Let $x \in X$ and $r > 0$. We will find a point in $b_r(x) \cap \bigcap_{n\geq 1} U_n$.

Since $\overline{A_1}$ has no interior, $V_1 := b_r(x) \cap U_1$ is non-empty and open. Thus there is a point $x_1 \in X$ and $0 < r_1 < r/2$ so that $\overline{b_{r_1}(x_1)} \subset V_1$. Proceed recursively. At stage n, we will have $x_1, \dots, x_n$ and $r_1, \dots, r_n$ so that $r_i < r/2^i$ and

$$\overline{b_{r_i}(x_i)} \subset V_i = b_{r_{i-1}}(x_{i-1}) \cap U_i \quad \text{for} \quad 1 \leq i \leq n.$$

Set $V_{n+1} = b_{r_n}(x_n) \cap U_{n+1}$. Since A_{n+1} is nowhere dense, this is a non-empty open set. So we may find a point x_{n+1} and an $r_{n+1} < r_n/2$ so that $\overline{b_{r_{n+1}}(x_{n+1})} \subset V_{n+1}$. This completes the inductive step.

The balls $\overline{b_{r_n}(x_n)}$ form a decreasing nested sequence of closed sets. We claim that the sequence $(x_n)_{n\geq 1}$ is Cauchy. Indeed, if $N \leq m < n$, then x_n, x_m lie in $\overline{b_{r_N}(x_N)}$ and hence $d(x_n, x_m) \leq 2r_N < 2^{1-N}r$. Since X is complete, this sequence has a limit, say $x_0 = \lim\limits_{n\to\infty} x_n$. Hence x_0 belongs to $\bigcap_{n\geq 1} \overline{b_{r_n}(x_n)}$. Since $\overline{b_{r_n}(x_n)}$ is disjoint from $\overline{A_n}$, we have $x_0 \in \bigcap_{n\geq 1} U_n$. Moreover, $x_0 \in V_1 \subset b_r(x)$, so that $d(x, x_0) < r$. Therefore $\bigcap_{n\geq 1} U_n$ is dense in X. ■

The following important result is sometimes called the Uniform Boundedness Principle.

2.4.3. BANACH-STEINHAUS THEOREM. *Let X be a Banach space, let Y be a normed vector space, and let $\mathcal{A} \subset \mathcal{B}(X, Y)$. Suppose that for each $x \in A$, $\sup_{A\in\mathcal{A}} \|Ax\| = K_x < \infty$. Then $\sup_{A\in\mathcal{A}} \|A\| < \infty$.*

PROOF. Let $X_n = \{x \in X : K_x \leq n\}$. This is a closed set since if $x_k \in X_n$ and $x_k \to x$ and $A \in \mathcal{A}$, then $\|Ax\| = \lim_{k\to\infty} \|Ax_k\| \leq n$. By assumption, $X = \bigcup_{n\geq 1} X_n$. By the Baire Category Theorem, there is an n_0 so that A_{n_0} has interior, say $b_r(x_0) \subset A_{n_0}$. If $\|x\| < 1$, then $x_0 + rx \in A_{n_0}$. Hence

$$\|Ax\| = \frac{1}{r}\|A(x_0 + rx) - Ax_0\| \leq \frac{2n_0}{r}.$$

Therefore $\sup_{A\in\mathcal{A}} \|A\| \leq \frac{2n_0}{r}$. ■

2.4.4. COROLLARY. *Let X be a Banach space, let Y be a normed vector space, and let $\mathcal{A} \subset \mathcal{B}(X, Y)$ such that $\sup_{A\in\mathcal{A}} \|A\| = \infty$. Then $B = \{x \in X :$ $\mathcal{A}x$ is bounded$\}$ is first category.*

PROOF. $B = \bigcup_{n\geq 1} B_n$ where $B_n = \{x : \sup_{A\in\mathcal{A}} \|Ax\| \leq n\}$. As in the previous proof, B_n is closed. If some B_{n_0} has interior, then as in the previous proof, $\mathcal{A}$ is bounded. Since it isn't bounded, each B_n is nowhere dense. Hence B is first category. ■

2.4.5. COROLLARY. *If X and Y are Banach spaces, $T_n \in \mathcal{B}(X,Y)$ and $\lim_{n\to\infty} T_n x$ exists for every $x \in X$, then $Tx := \lim_{n\to\infty} T_n x$ belongs to $\mathcal{B}(X,Y)$ and $\|T\| \le \sup \|T_n\| < \infty$.*

PROOF. Apply the uniform boundedness principle to $\{T_n : n \ge 1\}$. Since convergent sequences are bounded, $\{T_n x : n \ge 1\}$ is bounded for each $x \in X$. Therefore $\sup \|T_n\| = M < \infty$. The map T is clearly linear and $\|T\| \le M$, so $T \in \mathcal{B}(X,Y)$. ■

Now we establish a second important result using the Baire Category Theorem.

2.4.6. OPEN MAPPING THEOREM. *Let X and Y be Banach spaces, and let $T \in \mathcal{B}(X,Y)$. If $TX = Y$, then T is open; i.e. if $U \subset X$ is open, then TU is open.*

PROOF. Observe that $Y = TX = \bigcup_{n\ge1} Tb_n(X) = \bigcup_{n\ge1} \overline{Tb_n(X)}$, where $b_r(X) = b_r(0_X)$. Since Y is complete, the Baire Category Theorem shows that there is some n so that $\overline{Tb_n(X)} = n\overline{Tb_1(X)}$ has interior. Thus $\overline{Tb_1(X)}$ has interior, so it contains $b_r(y_0)$ for some $y_0 \in Y$ and $r > 0$. Therefore

$$\overline{Tb_1(X)} = \tfrac{1}{2}\overline{Tb_1(X)} - \tfrac{1}{2}\overline{Tb_1(X)} \supset \tfrac{1}{2}b_r(y_0) - \tfrac{1}{2}b_r(y_0) = b_r(0_Y).$$

We claim that for $\varepsilon > 0$, $\overline{Tb_1(X)} \subset Tb_{1+\varepsilon}(X)$. Let $y \in \overline{Tb_1(X)}$. Pick $x_1 \in b_1(X)$ with $\|y - Tx_1\| < \frac{\varepsilon r}{2}$. Then $y_2 := y - Tx_1 \in \overline{Tb_{\varepsilon/2}(X)}$. So there is an $x_2 \in X$ with $\|x_2\| < \frac{\varepsilon}{2}$ so that $\|y - Tx_1 - Tx_2\| < \frac{\varepsilon r}{4}$. Recursively we find $x_n \in X$ with $\|x_n\| < 2^{1-n}\varepsilon$ so that $\big\|y - \sum_{k=1}^n Tx_k\big\| < 2^{-n}\varepsilon r$. Let $x = \sum_{k\ge1} x_k$. Then $\|x\| < 1 + \sum_{k\ge1} 2^{1-k}\varepsilon = 1 + \varepsilon$. By continuity, $Tx = y$. So $b_r(Y) \subset \overline{Tb_1(X)} \subset Tb_{1+\varepsilon}(X)$.

Now if U is open in X and $x \in U$, then $b_\rho(x) \subset U$ for some $\rho > 0$. Hence $TU \supset Tb_\rho(x) \supset b_{r\rho}(Tx)$. Therefore TU is open. ■

One of the most cited consequences of the Open Mapping Theorem (OMT) is the following.

2.4.7. BANACH ISOMORPHISM THEOREM. *Let X and Y be Banach spaces, and let $T \in \mathcal{B}(X,Y)$. If T is a bijection, then T is invertible.*

PROOF. Clearly T^{-1} is a well-defined linear map. We need to show that it is continuous. By the OMT, T is open. So $Tb_1(X) \supset b_r(Y)$ for some $r > 0$. Therefore $T^{-1}b_1(Y) \subset b_{1/r}(X)$. That is, $\|T^{-1}\| \le \frac{1}{r}$. ■

There are many other consequences of the Open Mapping Theorem, and we collect a few here.

2.4.8. COROLLARY. *Let X and Y be Banach spaces, and let $T \in \mathcal{B}(X, Y)$. If T maps X onto Y, then $Y \simeq X/\ker T$.*

PROOF. Since T is continuous, $\ker T$ is a closed subspace of X, and thus $X/\ker T$ is a Banach space. Define $S : X/\ker T \to Y$ by $S(x + \ker T) = Tx$. This is well defined, linear and bijective. Moreover if $\dot{x} = x + \ker T$,

$$\begin{aligned}\|S\dot{x}\| &= \inf\{\|Tv\| : v \in x + \ker T\} \\ &\leq \inf\{\|T\|\,\|v\| : v \in x + \ker T\} = \|T\|\,\|\dot{x}\|.\end{aligned}$$

So $\|S\| \leq \|T\| < \infty$. Thus S is continuous, and so is an isomorphism by the Banach Isomorphism Theorem. ■

2.4.9. COROLLARY. *Let X be a Banach space with two complete norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If there is a constant C so that $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$, then there is a constant C' so that $\|x\|_2 \leq C'\|x\|_1$.*

PROOF. Consider the identity map $\mathrm{id} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$. By hypothesis this is a continuous bijection. Hence the inverse map is continuous by the Banach Isomorphism Theorem. ■

2.4.10. DEFINITION. A linear map $T : D \subset X \to Y$, where D is a not necessarily closed subspace of X, is a *closed linear map* if the graph $\mathcal{G}(T) = \{(x, Tx) : x \in D\}$ is closed in $X \oplus_\infty Y$.

2.4.11. CLOSED GRAPH THEOREM. *Let X and Y be Banach spaces, and let $T : X \to Y$ be a linear map defined on all of X. If T is closed, then T is continuous.*

PROOF. Since $\mathcal{G}(T)$ is a closed subspace of $X \oplus_\infty Y$, it is a Banach space. The coordinate projections $\pi_1 : \mathcal{G}(T) \to X$ and $\pi_2 : \mathcal{G}(T) \to Y$ are continuous. Moreover π_1 is a bijection. By the Banach Isomorphism Theorem, π_1^{-1} is continuous. Therefore $T = \pi_1^{-1}\pi_2$ is continuous. ■

2.4.12. COROLLARY. *Let X and Y be Banach spaces, and let $T : X \to Y$ be a linear map defined on all of X. Then T is continuous if*

$$(\ddagger) \qquad x_n \in X,\ x_n \to 0 \ \text{and}\ Tx_n \to y \quad \Longrightarrow \quad y = 0.$$

PROOF. We show that condition (‡) implies that T is closed. Suppose that $(x_n, Tx_n) \in \mathcal{G}(T)$ converge to $(x_0, y_0) \in X \oplus_\infty Y$. Then $x_n - x_0 \to 0$ and $T(x_n - x_0) \to y_0 - Tx_0$. By (‡), $Tx_0 = y_0$, so that $(x_0, y_0) \in \mathcal{G}(T)$. Thus $\mathcal{G}(T)$ is closed. Hence T is continuous by the Closed Graph Theorem. ■

2.4.13. EXAMPLE. Hellinger-Toeplitz. Let T be a linear operator on a Hilbert space H such that $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$. Suppose that $x_n \in H$, $x_n \to 0$ and $Tx_n \to y_0$. Then

$$\|y_0\|^2 = \lim_{n\to\infty} \langle Tx_n, y_0\rangle = \lim_{n\to\infty} \langle x_n, Ty_0\rangle = 0.$$

Thus $y_0 = 0$. By the Closed Graph Theorem, T is continuous.

2.5. Fourier series

If $f \in L^1(\mathbb{T})$, the Fourier coefficients are

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{-in\theta}\,d\theta \quad \text{for} \quad n \in \mathbb{Z}.$$

It is easy to see that $|\hat{f}(n)| \le \|f\|_1$. The Riemann-Lebesgue Lemma states that $\lim_{|n|\to\infty} |\hat{f}(n)| = 0$. This follows from the density of the trigonometric polynomials. Hence $\Gamma f = \big(\hat{f}(n)\big)_{n\in\mathbb{Z}}$ belongs to $c_0(\mathbb{Z})$.

Define $s_n(f)(e^{it}) = \sum_{k=-n}^{n} \hat{f}(k)e^{ikt}$ be the partial sums of the Fourier series. Compute

$$\begin{aligned} s_n(f)(e^{it}) &= \sum_{k=-n}^{n} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{-ik\theta}\,d\theta e^{ikt} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta}) \sum_{k=-n}^{n} e^{ik(t-\theta)}\,d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})D_n(t-\theta)\,d\theta \end{aligned}$$

and we sum the geometric series to get

$$\begin{aligned} D_n(t) &= \sum_{k=-n}^{n} e^{ikt} = \frac{e^{i(n+1)t} - e^{-int}}{e^{it} - 1} = \frac{e^{i(n+\frac{1}{2})t} - e^{-i(n+\frac{1}{2})t}}{e^{it/2} - e^{-it/2}} \\ &= \frac{e^{i(n+\frac{1}{2})t} - e^{-i(n+\frac{1}{2})t}}{2i} \cdot \frac{2i}{e^{it/2} - e^{-it/2}} = \frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}. \end{aligned}$$

Now we apply this when $f \in C(\mathbb{T})$. It is a classical result that there are continuous functions such that the Fourier series does not converge at some points. In fact, this behaviour is typical.

2.5.1. THEOREM. *For $t_0 \in [-\pi, \pi]$, the set*

$$\{f \in C(\mathbb{T}) : s_n(f)(e^{it_0}) \to f(e^{it_0})\}$$

is first category.

PROOF. The map $\varphi_n(f) = s_n(f)(e^{it_0}) = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(e^{i\theta})D_n(t_0-\theta)\,d\theta$ is a continuous linear functional on $C(\mathbb{T})$. We will show that $\|\varphi_n\| = \|D_n\|_1 \to \infty$. Clearly if $\|f\|_\infty \le 1$, then

$$|\varphi_n(f)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t_0-\theta)|\,d\theta \le \|D_n\|_1.$$

For $\varepsilon > 0$, define

$$f_\varepsilon(t) = \begin{cases} 1 & \text{if } D_n(t_0-\theta) \ge \varepsilon \\ -1 & \text{if } D_n(t_0-\theta) \le -\varepsilon \\ \text{p.w.linear} & \text{in between} \end{cases}$$

Then $\|f_\varepsilon\|_\infty = 1$ and for ε sufficiently small, $\varphi_n(f_\varepsilon) \approx \|D_n\|_1$. Now we estimate this norm.

$$\begin{aligned}\|D_n\|_1 &= \int_{-\pi}^{\pi} |D_n(t)|\,dt = \frac{1}{\pi}\int_0^{\pi} \frac{|\sin(n+\frac12)t|}{\sin\frac12 t} \\ &\ge \frac{1}{\pi}\int_0^{\pi}\frac{|\sin(n+\frac12)t|}{t/2} = \frac{2}{\pi}\int_0^{(n+\frac12)\pi} |\sin x|\,\frac{dx}{x} \\ &= \frac{2}{\pi}\int_0^{\pi/2} \sin x\,\frac{dx}{x} + \sum_{k=1}^{2n} \frac{1}{k\pi/2}\,\frac{2}{\pi}\int_{k\pi/2}^{(k+1)\pi/2} |\sin x|\,dx \\ &\ge \frac{4}{\pi^2}\sum_{k=1}^{2n}\frac{1}{k} \approx \frac{4}{\pi^2}\ln 2n \to \infty.\end{aligned}$$

Therefore $\mathcal{A} = \{\varphi_n : n \ge 1\}$ is unbounded. By Corollary 2.4.4,

$$\{f \in C(\mathbb{T}) : s_n(f)(e^{it_0}) \text{ is bounded}\}$$

is first category. This contains any f whose Fourier series converges at e^{it_0}. ■

2.5.2. THEOREM. *The map $\Gamma : L^1(\mathbb{T}) \to c_0(\mathbb{Z})$ is continuous and one to one, but is not surjective.*

PROOF. Clearly Γ is linear, and $\|\Gamma f\|_\infty = \sup|\hat f(n)| \le \|f\|_1$, so Γ is norm 1. The Riemann-Lebesgue Lemma shows that the range of Γ lies in $c_0(\mathbb{Z})$.

For $f \in L^1(\mathbb{T})$, define $f_r(e^{i\theta}) = \sum\limits_{n\in\mathbb{Z}} \hat f(n) r^{|n|} e^{in\theta}$. This series converges absolutely and uniformly to a continuous function for $0 \le r < 1$. To see that Γ is injective, we quote a result from Fourier analysis: $f_r \to f$ in $L^1(\mathbb{T})$ as $r \to 1^-$. This shows that f can be recovered from its Fourier series, and thus Γ is one to one.

If Γ were onto, then the Banach Isomorphism Theorem would show that Γ^{-1} is continuous; i.e. there would be a constant C so that $\|f\|_1 \le C\|\Gamma f\|_\infty$. However we saw in the previous proof that $\|D_n\|_1 \to \infty$ while $\hat{D}_n(k) = \begin{cases} 1 & \text{if } |k| \le n \\ 0 & \text{if } |k| > n. \end{cases}$ So $\|\hat{D}_n\|_\infty = 1$ and $\|\Gamma \hat{D}_n\|_\infty = 1$. Thus Γ^{-1} is not bounded. Therefore Γ is not surjective. ∎

2.6. The Hahn-Banach Theorems

2.6.1. DEFINITION. Let X be a vector space. A function $p : X \to \mathbb{R}$ is *sublinear* if $p(tx) = tp(x)$ for $t \ge 0$ and $p(x+y) \le p(x) + p(y)$ for $x, y \in X$.

2.6.2. EXAMPLES.
(1) On any normed space, $p(x) = \|x\|$ is sublinear.
(2) If $0 \in U \subset X$ is open and convex, then $X = \bigcup_{n \ge 1} nU$. The *Minkowski functional* is

$$p_U(x) = \inf\{t > 0 : x \in tU\}.$$

Note that if $x \in sU$ and $y \in tU$, then $x + y \in (s+t)U$ by convexity of U. That is, $\frac{x}{s}$ and $\frac{y}{t}$ are in U, so $\frac{s}{s+t}\frac{x}{s} + \frac{t}{s+t}\frac{y}{t} = \frac{x+y}{s+t} \in U$. Positive homogeneity is clear. So p_U is sublinear.

We start with a general result showing that functionals dominated by a sublinear functional on a subspace can be extended to one on the whole space. It starts with a "soft" part using Zorn's Lemma, followed by a "hard" part where an explicit extension is constructed increasing the space by one dimension.

2.6.3. THEOREM (Hahn-Banach)**.** *Let X be a $\mathbb{R}$ vector space, and let M_0 be a subspace (no topology). Let p be a sublinear map on X. Suppose that $f_0 : M \to \mathbb{R}$ is linear and $f_0(x) \le p(x)$ for $x \in M_0$. Then there is a linear functional $f : X \to \mathbb{R}$ so that $f|_{M_0} = f_0$ and $f(x) \le p(x)$ for all $x \in X$.*

PROOF. Consider pairs (f, M) where $M_0 \subset M \subset X$ is a subspace and $f : M \to \mathbb{R}$ is a linear functional such that $f|_{M_0} = f_0$ and $f(x) \le p(x)$ for all $x \in M$. Say $f_1, M_1) \le (f_2, M_2)$ is $M_1 \subset M_2$ and $f_2|_{M_1} = f_1$. This is a partial order on the collection of all such pairs.

We will show that it is inductive. Suppose that $\{(f_\alpha, M_\alpha) : \alpha \in A\}$ is a chain. Let $M = \bigcup_{\alpha \in A} M_\alpha$ and set $f = \bigcup_{\alpha \in A} f_\alpha$. Then it is easy to see that $(f_\alpha, M_\alpha) \le (f, M)$ for all $\alpha \in A$. Hence by Zorn's Lemma, there exists a maximal element, say (f, M).

Suppose that $M \ne X$. Pick $x \in X \setminus M$. We will try to extend f to a functional f' on $M' = M + \mathbb{R}x$ by setting $f'(x) = a$. We obtain necessary and sufficient conditions on a:

$$f'(m+tx) = f(m) + ta \le p(m+tx) \quad \text{for all} \quad m \in M \text{ and } t \in \mathbb{R}.$$

This is true for $t = 0$. When $t > 0$, we set $m' = m/t$ and obtain

$$a \le \frac{p(m+tx) - f(m)}{t} = p(m'+x) - f(m') \quad \text{for all} \quad m' \in M.$$

When $t < 0$, we set $m' = m/|t|$ and obtain

$$\frac{f(m) - p(m+tx)}{|t|} = f(m') - p(m'-x) \le a \quad \text{for all} \quad m' \in M.$$

Therefore

$$\sup_{m \in M} f(m) - p(m-x) \le a \in \inf_{m \in M} p(m+x) - f(m).$$

Such a value a exists if and only if the LHS $\le$ RHS.

If LHS $>$ RHS, there are vectors $m_1, m_2 \in M$ so that

$$f(m_1) - p(m_1 - x) > p(m_2 + x) - f(m_2).$$

Rearranging yields

$$f(m_1 + m_2) > p(m_1 - x) + p(m_2 + x) \ge p(m_1 + m_2).$$

This contradicts the condition on f. So we must have LHS $\le$ RHS. Choosing any value of a satisfying the inequality yields an extension to (f', M'). This contradicts the maximality of (f, M). So we have $M = X$, which completes the proof. ■

2.6.4. HAHN-BANACH THEOREM. *Let X be a normed space and let M be a (not necessarily closed) subspace of X. Suppose that f_0 is a continuous linear functional on M. Then there is an $f \in X^*$ with $f|_M = f_0$ and $\|f\| = \|f_0\|$.*

PROOF. Real case. Let $p(x) = \|f_0\|\,\|x\|$. This is a sublinear functional, and $f_0(m) \le \|f_0\|\,\|m\| = p(m)$ for all $m \in M$. By Theorem 2.6.3, there is an extension $f : X \to \mathbb{R}$ so that $f|_M = f_0$ and

$$\pm f(x) = f(\pm x) \le p(\pm x) = \|f_0\|\,\|x\|.$$

Thus $|f(x)| \le \|f_0\|\,\|x\|$. So $\|f\| = \|f_0\|$.

Complex case. Consider X as a real Banach space. Let $g_0(x) = \operatorname{Re} f_0(x)$ for $x \in M$. Extend g_0 to a real linear functional g on X with $\|g\| = \|g_0\| \le \|f_0\|$. Define $f(x) = g(x) + ig(-ix)$. Then $f : X \to \mathbb{C}$ is real linear, and

$$f(ix) = g(ix) - ig(x) = i\big(g(x) + ig(-ix)\big) = if(x).$$

Therefore f is complex linear. Now if $x \in M$,

$$\begin{aligned} f(x) &= g(x) + ig(-ix) = \operatorname{Re} f_0(x) + i \operatorname{Re} f_0(-ix) \\ &= \operatorname{Re} f_0(x) + i \operatorname{Re} -if_0(x) = \operatorname{Re} f_0(x) + i \operatorname{Im} f_0(x) = f_0(x). \end{aligned}$$

Thus f extends f_0. Finally, if $f(x) = e^{i\theta}|f(x)|$, then

$$|f(x)| = f(e^{-i\theta}x) = g(e^{-i\theta}x) \leq \|g_0\| \, \|e^{-i\theta}x\| \leq \|f_0\| \, \|x\|.$$

Hence $\|f\| \leq \|f_0\|$. ■

2.6.5. COROLLARY. *Let X be a Banach space. If $x_0 \in X$, then there is an $f \in X^*$ with $\|f\| = 1$ and $f(x_0) = \|x_0\|$. Hence*

$$\sup_{\substack{\|f\|=1 \\ f\in X^*}} |f(x)| = \|x\| \qquad \textit{for all } x \in X.$$

PROOF. Define f_0 on $M = \mathbb{C}x_0$ by $f_0(\lambda x_0) = \lambda\|x_0\|$. Then $\|f_0\| = 1$ and $f_0(x_0) = \|x_0\|$. By Hahn-Banach, extend f_0 to some $f \in X^*$ with $\|f\| = \|f_0\|$. This proves the non-trivial part of the equality in the supremum. ■

2.6.6. COROLLARY. *If X be a Banach space, then X^* separates points of X.*

PROOF. If $x \neq y$, there is some $f \in X^*$ so that $f(x-y) = \|x-y\| > 0$. So $f(x) \neq f(x)$. ■

2.6.7. COROLLARY. *If X be a Banach space, then X imbeds isometrically into X^{**} by $\iota : x \to \hat{x}$ where $\hat{x}(f) := f(x)$.*

PROOF. Clearly ι is linear. Moreover

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

Hence this map is an isometry. ■

2.6.8. DEFINITION. A Banach space X is *reflexive* if $\iota X = X^{**}$.

2.6.9. REMARK. Hilbert spaces are reflexive by two applications of Theorem 2.3.6. The Riesz Theorem that $L^p(\mu)^* = L^q(\mu)$ for $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ shows that $L^p(\mu)$ is reflexive. Finite dimensional Banach spaces are reflexive because if V is n-dimensional, then so is V^* and V^{**}. So ιX is an n-dimensional subspace of X^{**}, and therefore they are equal.

The space c_0 is not reflexive since $c_0^* = l_1$ and $l_1^* = l_\infty$. If $L^1(\mu)$ is infinite dimensional, then it is not reflexive either. Similarly $C(X)$ is not reflexive if it is infinite dimensional.

We will often drop the ι and consider X as a canonical subspace of X^{**}.

2.6.10. COROLLARY. *If X be a Banach space, M is a closed subspace of X and $x_0 \in X \setminus M$, then there is a functional $f \in X^*$ with $\|f\| = 1$ such that $f(x_0) = \operatorname{dist}(x_0, M)$.*

PROOF. Let $Y = X/M$ and $\dot{x}_0 = x_0 + M \in Y$. By definition of the quotient norm, $\|\dot{x}_0\| = \operatorname{dist}(x_0, M)$. By Corollary 2.6.5, there is a functional $g \in Y^*$ with $\|g\| = 1$ and $g(\dot{x}_0) = \|\dot{x}_0\|$. Define $f = gQ$ where $Q : X \to Y$ is the quotient map. Then $\|f\| \le \|g\|\,\|Q\| = 1$ and $f(x_0) = g(\dot{x}_0) = \|\dot{x}_0\| = \operatorname{dist}(x_0, M)$. ■

2.6.11. DEFINITION. If X is a Banach space and $Y \subset X$, the *annihilator* of Y in X^* is $Y^\perp = \{f \in X^* : f|_Y = 0\}$. If $Z \subset X^*$, the *preannihilator* of Z in X is $Z_\perp = \{x \in X : \hat{x}|_Z = 0\} = Z^\perp \cap X$.

2.6.12. LEMMA. *If $Y \subset X$, then $Y^\perp$ is a weak-$*$ closed subspace of X^*. If $Z \subset X^*$, then $Z_\perp$ is a norm closed subspace of X. Moreover*

$$(Y^\perp)_\perp = \overline{\operatorname{span} Y} \quad \textit{and} \quad (Z_\perp)^\perp = \overline{\operatorname{span} Z}^{w^*}.$$

PROOF. The first two statements are straightforward. Suppose that $M = \overline{\operatorname{span} Y}$. Then $Y^\perp = M^\perp$. Suppose that $x \notin M$. Then by Corollary 2.6.10, there is an $f \in M^\perp$ so that $f(x) \neq 0$. Hence $x \notin (M^\perp)_\perp$. That is $(M^\perp)_\perp = M$. The last claim is similar. ■

2.6.13. PROPOSITION. *Let M be a closed subspace of a Banach space X. Then there are canonical isometric isomorphisms*

$$M^* \simeq X^*/M^\perp \quad \textit{and} \quad (X/M)^* \simeq M^\perp.$$

PROOF. Let $R : X^* \to M^*$ be the restriction map $Rf = f|_M$. Then $\ker R = M^\perp$. Thus there is a well-defined map $\tilde{R} : X^*/M^\perp \to M^*$ given by $\tilde{R}(f + M^\perp) = Rf$. The Hahn-Banach Theorem shows that if $h \in M^*$, then there is a functional $f \in X^*$ with $\|f\| = \|h\|$ and $f|_M = h$. Thus $h = Rf = \tilde{R}(f + M^\perp)$. Moreover $\|h\| \le \|f + M^\perp\| \le \|f\| = \|h\|$. Therefore the map $\tilde{R}$ is a surjective isometry; and hence an isometric isomorphism.

Let $Q : X \to X/M$ be the quotient map. Define a map $Q^* : (X/M)^* \to M^\perp$ by $Q^* f = f \circ Q$. Indeed, $Q^* f$ is a linear functional on X that annihilates M, and thus belongs to $M^\perp$. On the other hand, if $g \in M^\perp$, define $f(x + M) = g(x)$. This is easily seen to be well defined, and $Q^* f = g$. So Q^* is surjective. Now $\|Q^* f\| = \|fQ\| \le \|f\|\,\|Q\| = \|f\|$. On the other hand, suppose that $Q^* f = g$ and $1 > \|x + M\| = \inf_{m \in M} \|x + m\|$. Then there is some $y = x + m$ with $\|y\| < 1$. Thus $|f(x + M)| = |g(y)| < \|g\|$. Taking the supremum over all $x + M$ in the open unit ball yields $\|f\| \le \|g\|$. Therefore Q^* is a surjective isometry, and thus an isometric isomorphism. ■

2.6.14. PROPOSITION. *X is reflexive if and only if X^* is reflexive.*

PROOF. If X is reflexive, then $X^{***} = (X^{**})^* = X^*$. So X^* is reflexive.

If $X^{**} \neq \iota X$, pick $z \in X^{**} \setminus \iota X$. By the previous Corollary, there is a functional $f \in X^{***}$ such that $f|_X = 0$ and $f(z) \neq 0$. So $f \in X^{***} \setminus \iota X^*$. ∎

2.6.15. PROPOSITION. *If X is reflexive, then all closed subspaces and quotients of X are reflexive.*

PROOF. Let $Y \subset X$ be a closed subspace. Let $\Phi \in Y^{**}$. By Hahn-Banach, it has an extension of the same norm to an element $\tilde{\Phi} \in X^{**}$. Since X is reflexive, there is some $x \in X$ so that $\iota x|_{Y^*} = \tilde{\Phi}$. Thus $\tilde{\Phi}$ is τ_{X^*} continuous. Hence $\Phi = \tilde{\Phi}|_{Y^*}$ is τ_{Y^*} continuous; and thus belongs to Y. Therefore Y is reflexive.

Now $(X/Y)^* = Y^{\perp} \subset X^*$. By the first part, $Y^{\perp}$ is reflexive. Therefore X/Y is reflexive by Proposition 2.6.14. ∎

Here is one more result, which combines the Hahn-Banach Theorem with the Uniform Boundedness Principle.

2.6.16. PROPOSITION. *Let X be a Banach space. A subset $A \subset X$ is bounded if and only if $f(A)$ is bounded for each $f \in X^*$.*

PROOF. Clearly if A is bounded, then $f(A)$ is bounded for $f \in X^*$. Conversely, we identify A with $\iota A \subset X^{**}$, and consider it as a subset of $\mathcal{B}(X^*, \mathbb{F})$. The hypothesis is that it is bounded on each $f \in X^*$. By the Banach-Steinhaus Theorem, ιA is bounded. Since ι is isometric by Corollary 2.6.7, A is bounded. ∎

2.6.17. EXAMPLE. Banach Limits. There is a linear functional $L \in l_\infty(\mathbb{R})^*$ such that

(1) $\liminf x_n \leq L((x_n)) \leq \limsup x_n$, and

(2) If $y_{n+k} = x_n$ for $n \geq N$ and some $k \in \mathbb{Z}$, then $L((y_n)) = L((x_n))$.

This yields a generalized limit since $L((x_n)) = \lim x_n$ when the limit exists, and it is translation invariant: $L((x_{n+1})) = L((x_n))$.

PROOF. Let S be the shift map: if $x = (x_n)$, then $(Sx)_n = x_{n+1}$; i.e., $Sx = (x_{n+1})$. Let $M = \{x - Sx : x \in l_\infty(\mathbb{R})\}$. This is a subspace because S is linear.

Claim: $\operatorname{dist}(\mathbf{1}, M) = 1$, where $\mathbf{1} = (1, 1, 1, \dots)$. Indeed, $\|\mathbf{1} - 0\| = 1$. Suppose that $\|\mathbf{1} - (x - Sx)\| \leq r < 1$. Then $|1 - (x_n - x_{n+1})| \leq r$, so that $x_{n+1} \leq x_n - (1 - r)$. Recursively we get that $x_{n+1} \leq x_1 - n(1 - r)$. So (x_n) is not bounded, a contradiction. By Corollary 2.6.10, there is a linear functional $L \in l_\infty(\mathbb{R})^*$ with $1 = \|L\| = L(\mathbf{1})$ and $L|_M = 0$.

Claim: $c_0 \subset \overline{M} \subset \ker L$. If $x \in c_0$, then

$$x - S^n x = (x - Sx) + (Sx - S^2x) + \cdots + (S^{n-1}x - S^n x) \in M \quad \text{for} \quad n \geq 1.$$

Moreover $\lim_{n\to\infty} S^n x = 0$. Therefore $x \in \overline{M}$.

Thus $L(S^k x) = L(x)$ and the initial terms of the sequence are irrelevant to $L(x)$, so (2) holds. Let $x \in l_\infty(\mathbb{R})$ and let $\alpha = \liminf x_n$ and $\beta = \limsup x_n$. If $\alpha = \beta$, i.e. if $\lim_{n\to\infty} x_n = \alpha$, then $x - \alpha\mathbf{1} \in c_0$; so that $L(x) = \alpha$. If $\alpha < \beta$, let $y = \frac{2}{\beta-\alpha}\big(x - \frac{\alpha+\beta}{2}\mathbf{1}\big)$; so that $\liminf y_n = -1$ and $\limsup y_n = 1$. Then there is a $z \in c_0$ so that $\|y - z\|_\infty = 1$. Thus $L(y) = L(y - z) \in [-1, 1]$. Rescaling, we get that $\alpha \leq L(x) \leq \beta$. ■

If we want a functional on complex l_∞, we just complexify: $\tilde{L}(x + iy) := L(x) + iL(y)$. You can check that $\|\tilde{L}\| = 1$ and that it is translation invariant, kills c_0, and yields the limit value when it exists.

2.6.18. EXAMPLE. Runge's Theorem. Let $K \subset \mathbb{C}$ be a compact set such that $\mathbb{C} \setminus K$ is connected. Then every function $f(z)$ which is analytic on an open neighbourhood $U \supset K$ is a uniform limit of polynomials on K.

PROOF. Let $P(K)$ be the closure of the polynomials in $C(K)$ and let $R(K)$ be the closure of the rational functions with poles in K^c. There is a union $\mathcal{C}$ of finitely many closed curves in $U \setminus K$ so that the winding number of $\mathcal{C}$ satisfies $\operatorname{ind}_{\mathcal{C}}(z) = \begin{cases} 1 & \text{if } z \in K \\ 0 & \text{if } z \in U^c \end{cases}$. Therefore by Cauchy's Theorem, for $z \in K$,

$$f(z) = \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{f(w)}{w - z}\,dw \approx \frac{1}{2\pi i}\sum \frac{f(w_k)}{w_k - z}(w_k - w_{k-1})$$

where the last term is a Riemann sum that uniformly approximates the integral for $z \in K$. It follows that $f|_K$ lies in the closure, $\overline{\operatorname{span}\{\frac{1}{a-z}\big|_K : a \in K^c\}} \subset R(K)$.

Suppose that $|a| > M := \sup\{|z| : z \in K\}$. Then

$$\frac{1}{a - z} = \frac{1}{a}\frac{1}{1 - z/a} = \frac{1}{a}\sum_{n\geq 0}\big(\frac{z}{a}\big)^n.$$

This series converges uniformly on K, and thus $\frac{1}{a-z}\big|_K \in P(K)$.

Now let $\varphi \in C(K)^*$ such that $\ker\varphi \supset P(K)$. Define $h(a) = \varphi(\frac{1}{w-z})$ for $a \in \mathbb{C} \setminus K$. Claim: h is analytic on $\mathbb{C} \setminus K$. Suppose that $b_{2r}(a) \subset \mathbb{C} \setminus K$. Then for $|w| < r$,

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

This converges uniformly for $z \in K$ because $|w| \le r < 2r \le |z-a|$. Hence we can apply φ to get

$$h(a+w) = \sum_{n\ge 0} \varphi\Big(\frac{-1}{(z-a)^{n+1}}\Big) w^n.$$

Since this is valid for $|w| < r$, h is analytic near a. Now h vanishes on $\{a : |a| > M\}$, so $h = 0$ on the connected component containing this annulus, which is $\mathbb{C} \setminus K$ by hypothesis. Hence $\ker \varphi \supset \overline{\operatorname{span}\{\frac{1}{a-z}\big|_K : a \in K^c\}}$.

We deduce that $\varphi(f) = 0$. However if $f \notin P(K)$, there would be a functional φ annihilating $P(K)$ but non-zero on f. Therefore $f \in P(K)$. In particular, this applies to any rational function with poles off of K. So $R(K) = P(K)$. ■

2.7. Complemented Subspaces

2.7.1. DEFINITION. Let X be a Banach space. A closed subspace $Y \subset X$ is *complemented* if there is a closed subspace $Z \subset X$ so that $Y \cap Z = \{0\}$ and $Y + Z = X$. This is called an *internal direct sum*; and we write $X = Y \dotplus Z$.

Define $T : Y \oplus_1 Z \to X$ by $T(y,z) = y + z$. By assumption, T is a bijection. Indeed, if $T(y,z) = 0$, then $y = -z \in Y \cap Z = \{0\}$, so $(y,z) = (0,0)$. Therefore T is one to one. It is onto by the second condition. Also

$$\|y+z\| \le \|y\| + \|z\| = \|(y,z)\|.$$

So $\|T\| \le 1$. By Banach's Isomorphism Theorem, T is an isomorphism; i.e. T^{-1} is continuous. Thus $X \simeq Y \oplus Z$. In particular, $Z \simeq X/Y$.

Define $P : Y \oplus_1 Z \to Y \oplus_1 0$ by $P(y,z) = (y,0)$. This is idempotent: $P^2 = P$ and $\|P\| = 1$. Then TPT^{-1} is a continuous projection of X onto Y with kernel Z. It is easy to see that this is the unique idempotent with this property.

Conversely suppose that $Q \in \mathcal{B}(X)$ is a continuous idempotent with range Y; and let $Z = \ker Q = \operatorname{Ran}(I-Q)$. Then $X \ni x = Qx + (I-Q)x \in Y + Z$; and if $x \in Y \cap Z$, then $x = (I-Q)x' = Qx = Q(I-Q)x' = 0$. So Y is complemented.

2.7.2. EXAMPLE. X^* is always complemented in X^{***}. Let $\iota : X^* \to X^{***}$ be the canonical injection, and let $R : X^{***} \to X^*$ be the restriction map $(R\varphi) = \varphi|_X$. Note that $R\iota = I_{X^*}$. Then $P = \iota R$ is a continuous projection of X^{***} onto X^* because $P^2 = \iota(R\iota)R = P$.

Now we will show that a familiar subspace of a Banach space is not complemented.

2.7.3. LEMMA. *There are uncountably many subsets $\{A_r \subset \mathbb{N} : r \in \mathbb{R}\}$ such that $A_r \cap A_s$ is a finite set if $r \neq s$.*

PROOF. List $\mathbb{Q} = \{q_n : n \geq 1\}$. For each $r \in \mathbb{R}$, pick a sequence $(q_{n(r,i)})_{i\geq 1}$ of rational numbers with r as its limit. Let $A_r = \{n(r,i) : i \geq 1\}$. Sequences converging to distinct points can only have finitely many terms in common. ■

2.7.4. THEOREM. *c_0 is not complemented in l_∞.*

PROOF. Suppose that $l_\infty \simeq c_0 \oplus Z$ where $Z \simeq l_\infty/c_0$. Observe that if $z = (z_n) + c_0 \in Z$, then $\|z\| = \limsup |z_n|$. Let A_r be the sets constructed in Lemma 2.7.3, and let $x_r = \chi_{A_r} + c_0 \in l_\infty/c_0$. Claim: if $r_1, \dots, r_n$ are distinct, $\big\| \sum_{i=1}^n a_i x_{r_i} \big\| = \max\{|a_i| : 1 \leq i \leq n\}$. Let $B_i = A_{r_i} \setminus \bigcup_{j\neq i} A_{j_j}$; so $A_{r_i} \setminus B_i$ is finite and $B_1, \dots, B_n$ are disjoint and infinite. Thus $x_{r_i} = \chi_{B_i} + c_0$. Hence

$$\Big\| \sum_{i=1}^n a_i x_{r_i} \Big\| \leq \Big\| \sum_{i=1}^n a_i \chi_{B_i} + c_0 \Big\| = \max\{|a_i| : 1 \leq i \leq n\}.$$

Now suppose that $T : l_\infty/c_0 \to l_\infty$ is a continuous one to one map into l_∞. Let $y_r = Tx_r \neq 0$; so there is an integer n_r so that $y_r(n_r) \neq 0$. Now

$$\mathbb{R} = \bigcup_{n,k\in\mathbb{N}} \{r : |y_r(n)| > 1/k\}.$$

Since this is a countable union, there is some n, k so that $\{r : |y_r(n)| > 1/k\}$ is uncountable. Let r_j for $j \geq 1$ be distinct reals in this set. Write $y_{r_j}(n) = e^{i\theta_j}|y_{r_j}(n)|$. Then $\big\| \sum_{j=1}^N e^{-i\theta_j} x_{r_i} \big\| = \max\{|e^{-i\theta_j}|\} = 1$, but

$$\Big\| T \sum_{j=1}^N e^{-i\theta_j} x_{r_i} \Big\| \geq \Big| \sum_{j=1}^N e^{-i\theta_j} y_{r_i}(n) \Big| \geq \frac{N}{k}.$$

This tends to ∞ as $N \to \infty$, and thus T is not bounded. This contradiction shows that c_0 is not complemented in l_∞. ■

We can parlay this result into a more general fact.

2.7.5. THEOREM. *c_0 is not isomorphic to a complemented subspace of any dual space X^*.*

PROOF. Suppose that $X^* \simeq c_0 \oplus_1 Z$ for some complementary Banach space Z. Then $X^{***} \simeq l_\infty \oplus_1 Z^{**}$. Let $J : c_0 \to X^*$ be a continuous injection onto a closed subspace of X, and let Q be a projection of X^* onto Jc_0. Let $R : X^{***} \to X^*$ be the restriction map to X. Define

$$P = J^{-1}QRJ^{**} : l_\infty \xrightarrow{J^{**}} X^{***} \xrightarrow{R} X^* \xrightarrow{J^{-1}Q} c_0 \xrightarrow{\iota} l_\infty.$$

Then if $x \in c_0$,

$$Px = \iota J^{-1}QRJx = \iota J^{-1}QJx = \iota J^{-1}Jx = \iota x$$

because $Jc_0 \subset X^*$ and $R|_{X^*} = I_{X^*}$ and $Jx = QJx$. Therefore $P^2 = P$, and this is a continuous projection of l_∞ onto c_0, which is impossible by Theorem 2.7.4. ■

2.7.6. COROLLARY. *If (K, d) is an infinite compact metric space, then $C(K)$ is not a dual space.*

PROOF. Pick a convergent sequence of distinct points in K, say $(x_i)_{i\geq 1}$ with $\lim_{i\to\infty} x_i = x_0$. Define $r_i = \min\{d(x_i, x_j) : j \neq i\}$. Then $r_i > 0$. Define $g_i(x) = \frac{1}{r_i}\max\{0, r_i - d(x, x_i)\}$. Then these are positive functions of norm 1 with disjoint supports: $g_i g_j = 0$ if $i \neq j$. Define $J : c_0 \to C(K)$ by $j((a_i)) = \sum_{i\geq 1} a_i g_i$. Because $a_i \to 0$, this sum converges uniformly, and thus belongs to $C(K)$ and $Ja(x_i) = a_i$ and $Ja(x_0) = 0$. Moreover $\|Ja\|_\infty = \|a\|_\infty$; i.e, J is an isometry. In particular it has closed range, say Y.

Define $T : C(K) \to I(x_0) = \{f : f(x_0) = 0\}$ by $TF = f - f(x_0)$. Then define $S : I(x_0) \to c_0$ by $Sf = (f(x_i))_{i\geq 1}$. Note that $STJa = SJa = a$ for $a \in c_0$ Let $P = JST$ and note that $P^2 = J(STJ)ST = JST = P$ is a continuous projection onto Y.

Since c_0 is isomorphic to a complemented subspace of $C(K)$, Theorem 2.7.5 shows that $C(K)$ is not a dual space. ■

Exercises for Chapter 2

1. (a) Prove that no infinite dimensional Banach space has a countable basis *as a vector space* (i.e. a collection $\{e_n : n \geq 1\}$ so that every vector in X is a *finite* linear combination of $\{e_n : n \geq 1\}$).

(b) Let X be an infinite dimensional Banach space. Recursively construct a sequence of unit vectors x_n so that $\text{dist}(x_n, \text{span}\{x_i : i < n\}) > 1 - 2^{-n}$. Hence deduce that the unit ball of X is not compact.

2. (a) Prove that $l_1^* = l_\infty$.

(b) Describe all infinite matrices $T = \big[t_{ij}\big]_{i,j=1}^\infty$ which act as bounded operators from l_1 to itself and find a formula for $\|T\|$.

3. Recall that $L_a^2(\mathbb{D})$ is the Hilbert space of analytic functions on the open unit disc which are square integrable with respect to area measure. Let $\zeta_n(z) = z^n$ for $n \geq 0$. Find constants α_n so that $\{\alpha_n\zeta_n : n \geq 0\}$ is an orthonormal basis for $L_a^2(\mathbb{D})$. Use this basis to find vectors $k_w \in L_a^2(\mathbb{D})$ so that $\langle h, k_w\rangle = h(w)$ for each $z \in \mathbb{D}$ and $h \in L_a^2(\mathbb{D})$. Be sure to sum the series and express k_w as a closed form function.

4. Let H be a Hilbert space which is a vector space of functions on a set X endowed with an inner product with the property that the linear functionals $\varepsilon_x(f) = f(x)$ are continuous for each $x \in X$.
(a) Show that for each $x \in X$, there is an element $k_x \in H$ so that $\langle f, k_x \rangle = f(x)$ for $x \in H$. Prove that the closed linear span of $\{k_x : x \in X\}$ is H.
(b) A multiplier of H is a function h on X so that $hf \in H$ for every $f \in H$. Prove that the linear map $M_h f = hf$ is continuous.

5. A *basis* for a Banach space X is a sequence $\{e_n : n \geq 1\}$ such that for each $x \in X$, there are unique scalars $\{c_n\}$ such that $x = \lim_{n\to\infty} \sum_{i=1}^{n} c_i e_i$. For convenience, normalize so that $\|e_n\| = 1$ for $n \geq 1$.
(a) Show that $\varphi_n(x) = c_n$ is a linear functional.
(b) Define $S_n x = \sum_{i=1}^{n} c_i e_i$, and set $|||x||| = \sup_{n\geq 1} \|S_n x\|$. Prove that $|||\cdot|||$ is a norm.
(c) Show that $(X, |||\cdot|||)$ is complete.
(d) Prove that the identity map T from $(X, |||\cdot|||)$ to $(X\| \cdot \|)$ is an isomorphism. Hence deduce that $\sup_{n\geq 1} \|S_n\| = \|T^{-1}\|$ and that each φ_n is continuous.

6. Let $\varphi(f) = \displaystyle\int_0^1 f(t)\,dt$ for $f \in C_{\mathbb{R}}[0,1]$. Let Φ be any Hahn–Banach extension of φ to the Banach space $B_{\mathbb{R}}[0,1]$ of all bounded real-valued functions on $[0,1]$ with the sup norm.
(a) What are the possible values for $\Phi(\chi_{[0,.5]})$?
(b) What are the possible values for $\Phi(\chi_{\mathbb{Q}\cap[0,1]})$?

7. Let X be a *separable* Banach space.
(a) Show that X is isometrically isomorphic to a subspace of l_∞.
HINT: find a countable set of linear functionals $\{\varphi_n\}$ of norm one so that $\sup |\varphi_n(x)| = \|x\|$ for all $x \in X$.
(b) Show that X is isometrically isomorphic to a quotient of l_1.
HINT: define a norm one linear map of l_1 *onto* X so that the image of the unit ball of l_1 is dense in the ball of X. Prove that the quotient norm coincides with the norm on X.

8. (a) Let X be a Banach space. If Y is a weak-$*$ closed subspace of X^*, let $Y_\perp = \{x \in X : f(x) = 0 \text{ for all } f \in Y\}$. Show that $Y = (Y_\perp)^\perp$.
(b) Hence show that Y is a dual space; and that the weak-$*$ topology on Y as this dual space coincides with the weak-$*$ topology of X^* restricted to Y.

CHAPTER 3

LCTVSs and Weak Topologies

3.1. Locally convex topological vector spaces

3.1.1. DEFINITION. A *seminorm* on a vector space V is a function $p : V \to [0, \infty)$ which is positive homogeneous and satisfies the triangle inequality, but non-zero vectors of norm 0 are allowed.

3.1.2. DEFINITION. A *locally convex topological vector space* (LCTVS) is a vector space X over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ together with a family $\mathcal{P}$ of seminorms on X such that $p(x) = 0$ for all $p \in \mathcal{P}$ implies that $x = 0$. Put a topology $\tau_{\mathcal{P}}$ on X generated by the convex sets

$$U(x_0, p, r) = \{x \in X : p(x - x_0) < r\} \quad \text{for} \quad x_0 \in X, p \in \mathcal{P} \text{ and } r > 0.$$

3.1.3. THEOREM. *Let X be a LCTVS.*

(1) *A neighbourhood base at 0 is provided by the convex open sets*

$$U_{F,r} = \{x \in X : p(x) < r,\ p \in F\} \quad \textit{for} \quad F \subset \mathcal{P} \textit{ and } r > 0.$$

The neighbourhood base at x_0 has the form $x_0 + U_{F,r}$.

(2) *X is Hausdorff.*

(3) *$X \times X \ni (x, y) \to x + y \in X$ is continuous.*

(4) *$\mathbb{F} \times X \ni (\alpha, x) \to \alpha x \in X$ is continuous.*

(5) *A net $(x_\lambda)_\Lambda$ in X converges to x if and only if $p(x - x_\lambda) \to 0$ for all $p \in \mathcal{P}$.*

PROOF. Observe that $\tau_{\mathcal{P}}$ has a base obtained from finite intersections of sets $U(x_i, p_i, r_i)$ for $1 \le i \le n$. Suppose that $0 \in U = \bigcap_{i=1}^n U(x_i, p_i, r_i)$. Let $r = \min_{1 \le i \le n} r_i - p_i(x_i - 0)$ and $F = \{p_1, \dots, p_n\}$. Then $0 \in U_{F,r} \subset U$ because if $x \in U_{F,r}$, then $p_i(x_i - x) \le p_i(x_i) + p_i(x) < p_i(x_i) + r_i - p(x_i) = r_i$; whence $x \in U$. So the sets $U_{F,r}$ provide a neighbourhood base around 0. The topology is translation invariant because $U(x_0, p, r) = x_0 + U(0, p, r)$.

If $x \neq y$, then $x - y \neq 0$ so there is some $p \in \mathcal{P}$ so that $p(x - y) = r > 0$. Hence $U(x, p, r/2)$ and $U(y, p, r/2)$ are disjoint open sets separating x and y.

Suppose that $x_0, y_0 \in X$ and $x_0 + y_0 \in U$ is open. Let $A(x,y) = x + y$ be the addition map. Choose $F \subset \mathcal{P}$ and $r > 0$ so that $x_0 + y_0 + U_{F,r} \subset U$. Claim: $(x_0 + U_{F,r/2}) \times (y_0 + U_{F,r/2}) \subset A^{-1}(x_0 + y_0 + U_{F,r})$. Indeed, if $x \in x_0 + U_{F,r/2}$ and $y \in y_0 + U_{F,r/2}$, then for $p \in F$,

$$p(x_0 + y_0 - x + y) \le p(x_0 - x) + p(y_0 - y) < \frac{r}{2} + \frac{r}{2} = r.$$

Hence A is continuous. The proof that scalar multiplication is continuous is similar.

Observe that $\lim_\Lambda x_\lambda = x$ means that for all finite $F \subset \mathcal{P}$ and $r > 0$, there is a $\lambda_0 \in \Lambda$ so that $x_\lambda \in x + U_{F,r}$ for all $\lambda \ge \lambda_0$. Taking $F = \{p\}$ shows that $\lim_\Lambda p(x - x_\lambda) = 0$ for all $p \in \mathcal{P}$. Conversely, when this holds, given $F = \{p_1, \dots, p_n\}$ and r, we can choose λ_i so that $p_i(x - x_\lambda) < r$ for $\lambda \ge \lambda_i$. Select λ_0 so that $\lambda_0 \ge \lambda_i$. Then for $\lambda \ge \lambda_0$, $x_\lambda \in x + U_{F,r}$. ■

3.1.4. EXAMPLES.

(1) Any normed space $(X, \|\cdot\|)$ is a LCTVS.

(2) Let X be a Banach space, and let $Y \subset X^*$ be a (not necessarily closed) subspace which separates points of X. Define seminorms $p_f(x) = |f(x)|$ for $f \in Y$. These seminorms determine the topology (X, τ_Y). Note that since Y is closed under scalar multiplication, $U_{F,r} = U_{\frac{1}{r}U,1}$. So the open sets $U_F := U_{F,1}$ form a base of neighbourhoods of 0.

If Z is a topological space, a map $T : Z \to (X, \tau_Y)$ is continuous if and only if $f \circ T$ is continuous for all $f \in Y$. This follows from Corollary 1.7.4.

Important special cases are (X, τ_{X^*}), the *weak topology* on X, and (X^*, τ_X), the *weak-$*$ topology* on X^*.

(3) Two other important examples are topologies on $\mathcal{B}(X,Y)$ where X, Y are Banach spaces. The *weak operator topology* is $(\mathcal{B}(X,Y), \text{WOT}) = (\mathcal{B}(X,Y), \tau_Z)$ where Z is the set of linear functions

$$f_{x,\varphi}(T) = \varphi(Tx) \quad \text{for} \quad x \in X \text{ and } \varphi \in Y^*.$$

Then $T_\lambda \xrightarrow{\text{WOT}} T$ if $\varphi(T_\lambda x) \to \varphi(Tx)$ for all $x \in X$ and $\varphi \in Y^*$.

The *strong operator topology* on $\mathcal{B}(X,Y)$ is $(\mathcal{B}(X,Y), \text{SOT})$ given by the seminorms $p_x(T) = \|Tx\|$ for $x \in X$. Then $T_\lambda \xrightarrow{\text{SOT}} T$ if $T_\lambda x \to Tx$ for all $x \in X$. This is the topology of pointwise convergence.

(4) Fréchet spaces. Let X be a LCTVS given by a *countable* family of seminorms $\{p_i : i \ge 1\}$. Then there is a translation invariant metric on X which yields the same topology. Define

$$d(x,y) = \sum_{i \ge 1} 2^{-i} \frac{p_i(x-y)}{1 + p(x-y)} \quad \text{for} \quad x, y \in X.$$

Since this depends only on $x-y$, it is translation invariant. The function $f(t) = \frac{t}{1+t}$ is increasing. Therefore

$$\frac{p(x+y)}{1+p(x+y)} \le \frac{p(x)+p(y)}{1+p(x)+p(y)} \le \frac{p(x)}{1+p(x)} + \frac{p(y)}{1+p(y)}.$$

This is valid for each p_i, and so d satisfies the triangle inequality. Since $d(x,y) = 0$ only if $p_i(x-y) = 0$ for all $i \ge 1$, this implies that $x = y$. Hence d is a metric. Now $x_\lambda \to x$ if and only if $p_i(x-x_\lambda) \to 0$ for $i \ge 1$ if and only if $\frac{p_i(x-x_\lambda)}{1+p(x-x_\lambda)} \to 0$ for $i \ge 1$. If $d(x, x_\lambda) \to 0$, then clearly $\frac{p_i(x-x_\lambda)}{1+p(x-x_\lambda)} \to 0$ for $i \ge 1$. Conversely, if $\frac{p_i(x-x_\lambda)}{1+p(x-x_\lambda)} \to 0$ for $i \ge 1$ and $\varepsilon > 0$, choose N so that $2^{-N} < \varepsilon/2$. Then choose $\lambda_0 \in \Lambda$ so that $\frac{p_i(x-x_\lambda)}{1+p(x-x_\lambda)} < \frac{\varepsilon}{2}$ for $1 \le i \le N$ and $\lambda \ge \lambda_0$. Then

$$d(x_\lambda, x) < \sum_{i=1}^{N} \frac{\varepsilon}{2} 2^{-i} + \sum_{i>N} 2^{-i} < \frac{\varepsilon}{2} + 2^N < \varepsilon.$$

Therefore the topology on X is equivalent to the metric topology given by d.

We say X is a *Fréchet space* if it is complete in this metric. It is an easy exercise to show that a sequence is Cauchy in this metric if and only if it is Cauchy in each seminorm p_i for $i \ge 1$.

Let $\Omega \subset \mathbb{C}$ be an open subset of the complex plane. Let $H(\Omega)$ be the space of analytic functions on Ω. For each compact $K \subset \Omega$, define $p_K(f) = \sup_{z\in K} |f(z)|$. Then $\mathcal{P} = \{p_K : K \subset \Omega,\ \text{compact}\}$ is a family of seminorms that makes $H(\Omega)$ a LCTVS. Observe that $f_\lambda \to f$ in $H(\Omega)$ if and only if f_λ converges to f uniformly on compact subsets of Ω, often called u.c.c. convergence. It is straightforward to show that the u.c.c. limit of analytic functions is analytic. It is not necessary to use all compact sets. Define $K_n = \{z \in \Omega : |z| \le n,\ \operatorname{dist}(z, \Omega^c) \ge \frac{1}{n}\}$. Then each K_n is compact, $K_n \subset K_{n+1}$, and $\bigcup_{n\ge1} \operatorname{int} K_n = \Omega$. Thus for any compact set $K \subset \Omega$, there is some n so that $K \subset K_n$. It follows that the seminorms p_{K_n} for $n \ge 1$ determine the topology. Since $H(\Omega)$ is complete in the topology of u.c.c. convergence, it is a Fréchet space.

3.1.5. PROPOSITION. *Let X be an LCTVS, and let $f : X \to \mathbb{F}$ be a linear map. The following are equivalent*

(1) *f is continuous.*

(2) *f is continuous at* 0*.*

(3) $\ker f$ *is closed.*

(4) *There is a finite subset $F = \{p_1, \dots, p_n\} \subset \mathcal{P}$ and $t_i \in \mathbb{R}^+$ so that $|f(x)| \le \sum_{i=1}^n t_i p_i(x)$ for all $x \in X$.*

PROOF. Clearly (1) implies (2). If (2) holds and $x_\lambda \in \ker f$ such that $x_\lambda \to x$, then $x - x_\lambda \to 0$. Thus $f(x) - f(x_\lambda) \to 0$ by continuity at 0, Since $f(x_\lambda) = 0$, we have $f(x) = 0$, so $\ker f$ is closed.

Suppose that (3) holds. We may assume that $f \neq 0$. Pick an $x_0 \in X$ so that $f(x_0) = 1$. Since $\ker f$ is closed, there is a neighbourhood $x_0 + U_{F,r}$ disjoint from $\ker f$. Since $x_0 + U_{F,r}$ is convex, so is $f(x_0 + U_{F,r}) = 1 + f(U_{F,r})$. Therefore $-1 \notin f(U_{F,r})$. However $U_{F,r}$ is *balanced*, meaning that if $x \in U_{F,r}$ and $|\lambda| \leq 1$ for $\lambda \in \mathbb{F}$, then $\lambda x \in U_{F,r}$. This is because $p_i(\lambda x) = |\lambda| p_i(x) \leq p_i(x) < r$ for $p_i \in F$. Therefore $f(U_{F,r}) \subset \mathbb{D} = \{z \in \mathbb{F} : |z| < 1\}$. It follows that if $\sum_{i=1}^n p_i(x) < r$, then $p_i(x) < r$ for $1 \leq i \leq n$, and so $|f(x)| < 1$. Scaling shows that $|f(x)| \leq \sum_{i=1}^n \frac{1}{r} p_i(x)$. So (4) holds.

Suppose that (4) holds: $|f(x)| \leq \sum_{i=1}^n t_i p_i(x)$ for all $x \in X$. Then if $x_\lambda \to x$, then $p_i(x - x_\lambda) \to 0$ and so $|f(x) - f(x_\lambda)| \leq \sum_{i=1}^n t_i p_i(x - x_\lambda) \to 0$. Therefore $f(x_\lambda) \to f(x)$, and f is continuous. ■

We will write X^* for the vector space of continuous linear functionals on a LCTVS X. We will not discuss the topology on X^*.

We now prove a useful lemma.

3.1.6. LEMMA. *Let $f_1, \ldots, f_n$ be linear functionals on a LCTVS X. Suppose that f is a linear functional on X such that $\ker f \supset \bigcap_{i=1}^n \ker f_i$. Then $f \in \operatorname{span}\{f_1, \ldots, f_n\}$.*

PROOF. We may assume that $f_1, \ldots, f_n$ are linearly independent because if say $f_n = \sum_{i=1}^{n-1} a_i f_i$, then $\ker f_n \supset \bigcap_{i=1}^{n-1} \ker f_i$. Thus there is no loss in dropping f_n from the list.

Define $T : X \to \mathbb{F}^n$ by $Tx = (f_1(x), \ldots, f_n(x))$. By linear independence, this map is surjective. Note that $\ker T = \bigcap_{i=1}^n \ker f_i$. So $X/\ker T \simeq \mathbb{F}^n$. Since $\ker f \supset \bigcap_{i=1}^n \ker f_i$, there is a well defined functional $\tilde{f}$ on $X/\ker T \simeq \mathbb{F}^n$ given by $\tilde{f}(x + \ker T) = f(x)$. Now every linear functional on $\mathbb{F}^n$ has the form $\varphi(y) = \sum_{i=1}^n a_i y_i$. It follows that there are scalars a_i so that $f(x) = \sum_{i=1}^n a_i f_i(x)$. Hence $f \in \operatorname{span}\{f_1, \ldots, f_n\}$. ■

3.1.7. COROLLARY. *A linear functional $f : (X, \tau_Y) \to \mathbb{F}$ is continuous if and only if $f \in Y$.*

PROOF. By definition, if $f \in Y$, then it is continuous. On the other hand, if f is continuous, then by Proposition 3.1.5, if f is continuous, then there are $f_1, \ldots, f_n \in Y$ so that $|f(x) \leq \sum_{i=1}^n t_i |f_i(x)|$. This forces $\ker f \supset \bigcap_{i=1}^n \ker f_i$. Hence by Lemma 3.1.6, f lies in $\operatorname{span}\{f_1, \ldots, f_n\} \subset Y$. ■

Under certain circumstances, convergent *sequences* must be bounded.

3.1.8. LEMMA. *Let X be a Banach space, and let Y be a closed subspace of X^* which norms X, i.e.* $\sup\{|f(x)| : f \in Y,\ \|f\| \le 1\} = \|x\|$. *Then if a sequence $x_n \xrightarrow{\tau_Y} x$, then $\{x_n : n \ge 1\}$ is bounded.*

PROOF. Consider $\hat{x}_n$ as linear functionals on Y by $\hat{x}_n(f) = f(x_n)$. The norming condition ensures that $\|\hat{x}_n\| = \|x_n\|$. Now $x_n \xrightarrow{\tau_Y} x$ means that $\hat{x}_n(f) \to \hat{x}(f)$ for all $f \in Y$. By the Banach-Steinhaus Theorem, $\{\hat{x}_n : n \ge 1\}$ is bounded. Therefore $\{x_n : n \ge 1\}$ is bounded. ∎

3.1.9. REMARK. This lemma applies in particular to the weak topology τ_{X^*} on X and to the weak-$*$ topology (X^*, τ_X).

If Y is not closed in X^*, then the continuous functionals on (X, τ_Y) is not complete. An example of this is the weak operator topology on $\mathcal{B}(X, Y)$ (Example 3.1.4) when X, Y are infinite dimensional. So Lemma 3.1.8 does not apply directly. However in this case, the conclusion is still valid. It is obtained using two applications of the Banach-Steinhaus Theorem using the completeness of X and Y separately.

3.1.10. EXAMPLE. Since $l_1 = c_0^*$, it has an associated weak-$*$ topology τ_{c_0}. Caveat: l_1 has many different preduals, and hence many different weak-$*$ topologies. Claim: a bounded net $(x_\lambda)_\Lambda$ converges τ_{c_0} to x_0 if and only if each coordinate $x_{\lambda,i} \to x_{0,i}$ for $i \ge 1$. One direction is clear since the element $\varepsilon_i \in c_0$ with a 1 in the ith coordinate and 0s elsewhere is in c_0; whence $x_{\lambda,i} = \varepsilon_i(x_\lambda) \to \varepsilon_i(x_0) = x_{0,i}$.

Conversely, suppose that $\sup_\Lambda \|x_\lambda\|_1 = M$ and that $x_{\lambda,i} \to x_{0,i}$ for $i \ge 1$. Let $y = (y_i) \in c_0$ and $\varepsilon > 0$. Pick K so that $\sup_{i>K} |y_i| < \varepsilon$. Then pick λ_0 so that for $\lambda \ge \lambda_0$, $\sum_{i=1}^K |x_{\lambda,i} - x_{0,i}|\,|y_i| < \varepsilon$, which is possible since the first K terms $|x_{\lambda,i} - x_{0,i}|$ each converge to 0. Then for $\lambda \ge \lambda_0$, since $\|x_\lambda - x_0\| \le M + \|x\|$,

$$\begin{aligned} |\langle x_\lambda - x_0, y\rangle| &\le \sum_{i=1}^K |x_{\lambda,i} - x_{0,i}|\,|y_i| + \sum_{i>K} |x_{\lambda,i} - x_{0,i}|\,|y_i| \\ &\le \varepsilon + (M + \|x\|)\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this shows that the net $(x_\lambda)_\Lambda$ converges in τ_{c_0} to x_0.

Next we construct an unbounded net that converges to 0 in τ_{c_0}. Let $\Lambda = \{F \subset c_0 : |F| < \infty\}$ ordered by inclusion. Use the Axiom of Choice to pick $x_F \in \bigcap_{y \in F} \ker y$ with $\|x_F\|_1 = |F|$. Since each $\ker y$ has codimension 1, this intersection has finite codimension, and hence is non-zero. This is clearly an unbounded net. However for any $y \in c_0$ and $F \subset \{y\}$, $\langle x_F, y\rangle = 0$. Hence $(x_F)_\Lambda \xrightarrow{\tau_{c_0}} 0$.

3.2. Geometric Hahn-Banach Theorem

A subspace $M \subset X$ has *finite codimension* if there are finitely many vectors $x_1, \dots, x_n$ such that $X = \operatorname{span}\{M, x_1, \dots, x_n\}$; and the codimension is the minimal number n which works. If $f \neq 0$ is a linear functional on X (not necessarily continuous), then pick some x_0 with $f(x_0) = 1$. Then we can write $x = (x - f(x)x_0) + f(x)x_0 \in \ker f + \mathbb{F}x_0$ for every $x \in X$. So $\ker f$ has codimension 1. By Proposition 3.1.5, $\ker f$ is closed precisely when it is continuous. So if f is discontinuous, $\overline{\ker f} = X$. It is easy to check that the intersection of finitely many subspaces of finite codimension is still finite codimension.

3.2.1. Definition. A set A is *affine* if $x, y \in A$, $t \in \mathbb{R}$, then $tx+(1-t)y \in A$. A *hyperplane* is an affine set of the form $H = \{x : \operatorname{Re} f(x) = a\}$ for f linear and $a \in \mathbb{R}$.

In the real case, H is a translate of a codimension 1 subspace. In the complex case, it is the translate of a *real codimension 1* subspace, $\ker \operatorname{Re} f$. In either case, H is closed if and only if f is continuous. A hyperplane splits the space X into two halves $H^+ = \{x : \operatorname{Re} f(x) > a\}$ and $H^- = \{x : \operatorname{Re} f(x) < a\}$.

3.2.2. Lemma. *Let X be a LCTVS, $0 \in U$ an open convex set, and let p_U be the Minkowski functional $p_U = \inf\{r > 0 : x \in rU\}$. Then p_U is continuous and $\{x : p_U(x) < 1\} = U$.*

Proof. Since $0 \in U$ and U is convex, $rU \subset sU$ if $0 < r < s$. Hence if $p_U(x) < 1$, then there is $r < 1$ so that $x \in rU \subset U$. Conversely, if $x \in U$, then since $(t, x) \to tx$ is continuous, there is some $\varepsilon > 0$ so that $(1+\varepsilon)x \in U$. Hence $p_U(x) \leq \frac{1}{1+\varepsilon} < 1$.

Suppose that $V \subset \mathbb{R}$ is open and $p_U(x_0) = t_0 \subset (t_0 - r, t_0 + r) \subset V$. If $x \in x_0 + r(U \cap -U)$, then $p_U(x - x_0) < r$ and $p_U(x_0 - x) < r$. By sublinearity,

$$p(x) \leq p(x - x_0) + p(x_0) < p(x_0) + r$$

and

$$p(x_0) \leq p(x_0 - x) + p(x) < p(x) + r.$$

Combining, we get $|p_U(x) - p_U(x_0)| < r$. Therefore $x_0 + r(U \cap -U) \subset p_U^{-1}(V)$. Hence p_U is continuous. ∎

3.2.3. Hyperplane Theorem. *Let X be a LCTVS and let $U \subset X$ be an open convex set such that $0 \notin U$. Then there is a continuous functional $f \in X^*$ so that $\operatorname{Re} f(x) > 0$ for all $x \in U$; i.e. $H = \{x : \operatorname{Re} f(x) = 0\}$ is a closed hyperplane through 0 which is disjoint from U.*

PROOF. Case: $\mathbb{F} = \mathbb{R}$. Fix $x_0 \in U$. Then $V = x_0 - U$ is an open convex neighbourhood of 0. By Lemma 3.2.2, p_V is continuous. Since $x_0 \notin V$, we have that $p_V(x_0) \geq 1$. Define $f_0(tx_0) = t$ for $t \in \mathbb{R}$. Then

$$f_0(tx_0) = \begin{cases} t \leq tp_V(x_0) = p_V(tx_0) & \text{if } t \geq 0 \\ t \leq 0 \leq p_V(tx_0) & \text{if } t < 0. \end{cases}$$

So $f_0 \leq p_V$ on $\mathbb{R}x_0$. By the Hahn-Banach Theorem 2.6.3, there is a linear functional f on X so that $f(x) \leq p_V(x)$ for all $x \in X$ and $f(x_0) = 1$. Note that

$$\begin{aligned} f^{-1}(b_1(0)) &= \{x : |f(x)| < 1\} = \{x : f(x) < 1\} \cap \{x : f(-x) < 1\} \\ &\supseteq \{x : p_V(x) < 1\} \cap \{x : p_V(-x) < 1\} = V \cap -V. \end{aligned}$$

Thus $f^{-1}(b_1(0))$ contains an open neighbourhood of 0. Therefore f is continuous. Finally if $x \in U$, then $x_0 - x \in V$, so that $f(x_0 - x) < p_V(x_0 - x) < 1$. Hence $f(x) > f(x_0) - 1 = 0$.

Case: $\mathbb{F} = \mathbb{C}$. Treat X as a real vector space and find a continuous real functional f using the real case. Set $g(x) = f(x) + if(-ix)$. As in the proof of the Hahn-Banach Theorem, g is a continuous, complex linear functional such that $\operatorname{Re} g = f$. Thus if $x \in U$, we have $\operatorname{Re} g(x) = f(x) > 0$. ∎

The following two special cases are immediate.

3.2.4. COROLLARY. *If X is a Banach space, $U \subset X$ is convex and open, and $0 \notin U$, then there is some $f \in X^*$ so that* $\operatorname{Re} f(x) > 0$ *for all $x \in U$.*

3.2.5. COROLLARY. *If X is a Banach space, $Y \subset X^*$ separates points, $U \subset X$ is convex and τ_Y-open, and* $0 \notin U$, *then there is some $f \in Y$ so that* $\operatorname{Re} f(x) > 0$ *for all $x \in U$.*

3.2.6. SEPARATION THEOREM. *Let X be a LCTVS and let A, B be disjoint convex subsets of X. If A is open, then there is a continuous linear functional $f \in X^*$ and $\alpha \in \mathbb{R}$ so that*

$$\operatorname{Re} f(b) \leq \alpha < \operatorname{Re} f(a) \quad \textit{for all} \quad a \in A,\ b \in B.$$

PROOF. Let $U = A - B = \{a - b : a \in A,\ b \in B\} = \bigcup_{b \in B} A - b$. This is open, convex and $0 \notin U$. By the Hyperplane Theorem, there is an $f \in X^*$ so that $\operatorname{Re} f(x) > 0$ for all $x \in U$. Thus for $a \in A$ and $b \in B$,

$$0 < \operatorname{Re} f(a - b) = \operatorname{Re} f(a) - \operatorname{Re} f(b).$$

Hence

$$\sup_{b \in B} \operatorname{Re} f(b) := \alpha \leq \inf_{a \in A} \operatorname{Re} f(a).$$

Now $f(A)$ is open. Indeed, suppose that $a \in U_{F,r}(a) \subset A$. If $U_{F,r}(0) \subset \ker f$, then $\ker f \supset \bigcup_{n\geq 1} U(F, nr)(0) = X$ and so $f = 0$. This isn't possible, and thus $f(U_{F,r}(0))$ contains some $z \neq 0$. Since $U_{F,r}(0)$ is balanced and convex, so is $f(U_{F,r}(0))$; and so $f(U_{F,r}(0))$ contains an open disc about 0. It follows that $f(A)$ is open, and thus the infimum $\inf_{a\in A} \operatorname{Re} f(a)$ is not attained. Therefore $\alpha < f(a)$ for all $a \in A$. ∎

The following is immediate.

3.2.7. COROLLARY. *Let X be a LCTVS and let A, B be disjoint open convex subsets of X. Then there is a continuous linear functional $f \in X^*$ and $\alpha \in \mathbb{R}$ so that*

$$\operatorname{Re} f(b) < \alpha < \operatorname{Re} f(a) \quad \textit{for all} \quad a \in A,\ b \in B.$$

3.2.8. LEMMA. *Let X be a LCTVS. Suppose that K is a compact subset of X, V is open in X and $K \subset V$. Then there is a convex open neighbourhood $U \ni 0$ such that $K + U \subset V$.*

PROOF. For each $x \in K$, there is a basic neighbourhood $x + U(F_x, r_x) \subset V$. The sets $\{x + U(F_x, \frac{r_x}{2}) : x \in K\}$ cover K, so there is a finite subcover $x_i + U(F_i, \frac{r_i}{2})$ for $1 \leq i \leq n$. Let $F = \bigcup_{i=1}^n F_i$ and $r = \min\{r_i : 1 \leq i \leq n\}$; and set $U = U(F, \frac{r}{2})$. If $x \in K$, then $x \in x_i + U(F_i, \frac{r_i}{2})$ for some i. Then

$$\begin{aligned} x + U &\subset x_i + U(F_i, \tfrac{r_i}{2}) + U(F, \tfrac{r}{2}) \\ &\subset x_i + U(F_i, \tfrac{r_i}{2}) + U(F_i, \tfrac{r_i}{2}) = x_i + U(F_i, r_i) \subset V. \end{aligned}$$

Hence $K + U \subset V$. ∎

3.2.9. COROLLARY. *Let X be a LCTVS and let A, B be disjoint closed convex subsets of X. If B is compact, then there is an $f \in X^*$ and $\beta < \alpha \in \mathbb{R}$ so that*

$$\sup_{b\in B} \operatorname{Re} f(b) := \beta < \alpha := \inf_{a\in A} \operatorname{Re} f(a).$$

PROOF. By Lemma 3.2.8, there is a convex open set $U \ni 0$ so that $B + U \subset A^c$. Then $B + U$ is convex and open, and disjoint from A. Hence there is $f \in X^*$ and $\alpha \in \mathbb{R}$ so that

$$\operatorname{Re} f(b) < \alpha = \inf_{a\in A} \operatorname{Re} f(a) \quad \text{for all} \quad b \in B + U.$$

Since B is compact and f is continuous, the value $\beta = \sup_{b\in B} \operatorname{Re} f(b)$ is attained, and thus $\beta < \alpha$. ∎

3.2.10. EXAMPLE. Consider (l_1, τ_{c_0}). Let $\{\delta_i : i \geq 1\}$ be the standard basis for l_1. Define

$$A = \left\{x = (x_n) \in l_1 : \sum_{n \geq 1} x_n = 0\right\} \quad \text{and} \quad B = \{\delta_1\}.$$

Clearly A and B are disjoint convex sets, and B is compact. However no functional $f \in c_0 = (l_1, \tau_{c_0})^*$ separates A from B. Indeed, suppose that $0 \neq f = (a_n) \in c_0$, say $a_m \neq 0$. Then $\delta_m - \delta_n \in A$, and $f(\delta_m - \delta_n) = a_m - a_n \to a_m \neq 0$. So the subspace A is not contained in $\ker f$, whence $f(A) = \mathbb{C}$. Therefore $f(\delta_1)$ cannot be disjoint from $f(A)$.

What went wrong is that A is not closed in the τ_{c_0} topology. Indeed, $A = \ker g$ where $g = (1, 1, 1, \dots) \in l_\infty = l_1^*$. By Proposition 3.1.5, A is not τ_{c_0} closed.

On the other hand, A is norm closed. The functional g separates A and B since $\sup_{a \in A} \operatorname{Re} g(a) = 0 < 1 = g(\delta_1)$.

Taking $A = \{x\}$ and $B = \{y\}$ yields:

3.2.11. COROLLARY. *Let X be a LCTVS. Then X^* separates points of X.*

3.2.12. DEFINITION. Let X be a LCTVS and let $A \subset X$. The convex hull of A, $\operatorname{conv}(A)$, consists of all points $x = \sum_{i=1}^n t_i x_i$ for $n \geq 1$, $x_i \in X$, and $t_i \geq 0$ such that $\sum_{i=1}^n t_i = 1$. Let $\overline{\operatorname{conv}}(A)$ denote the closure of $\operatorname{conv}(A)$.

3.2.13. COROLLARY. *Let X be a LCTVS and let $A \subset X$. Then the closed convex hull of A is the intersection of all closed half-spaces containing A:*

$$\overline{\operatorname{conv}}(A) = \bigcap \left\{ \{x : \operatorname{Re} f(x) \leq \alpha\} : f \in X^*, \ \sup_{a \in A} \operatorname{Re} f(x) = \alpha < \infty \right\}.$$

Hence if A is closed and convex, then $x \in A$ if and only if

$$\operatorname{Re} f(x) \leq \sup_{a \in A} \operatorname{Re} f(x) \quad \text{for all} \quad f \in X^*.$$

PROOF. The intersection of closed convex sets is closed and convex, and contains A by design. If $x \notin \overline{\operatorname{conv}}(A)$, then by Corollary 3.2.9, there is $f \in X^*$ so that $\operatorname{Re} f(x) > \alpha = \sup_{a \in A} \operatorname{Re} f(x)$. Thus the intersection is exactly $\overline{\operatorname{conv}}(A)$. ∎

3.2.14. PROPOSITION.

(1) *Every norm closed convex set in a Banach space X is weakly closed.*

(2) *Every norm closed ball in X^* is weak-$*$ closed.*

PROOF. A closed half space

$$\{x : \operatorname{Re} f(x) \leq \alpha\} = f^{-1}(\{z \in \mathbb{C} : \operatorname{Re} z \leq \alpha\})$$

for $f \in X^*$ is weakly closed. Thus a closed convex set is weakly closed by Corollary 3.2.13. The ball

$$\overline{b_r(x_0^*)} = \{x^* \in X^* : \|x^* - x_0^*\| \le r\} = \bigcap_{x \in \overline{b_1(X)}} \{x^* \in X^* : |\hat{x}(x^* - x_0^*)| \le r\}$$

is the intersection of weak-$*$ closed sets. ∎

3.2.15. GOLDSTINE'S THEOREM. *Let X be a Banach space. Then $b_1(X)$ is weak-$*$ dense in $b_1(X^{**})$. In particular, X is weak-$*$ dense in X^{**}.*

PROOF. Let $A = \overline{\iota b_1(X)}^{\tau_{X^*}}$. This set is τ_{X^*}-closed and convex. It is contained in $\overline{b_1(X^{**})}$ because the closed ball is also weak-$*$ closed by Proposition 3.2.14. If A is strictly smaller, let $x^{**} \in \overline{b_1(X^{**})} \setminus A$. By Corollary 3.2.9, there is some $f \in X^*$ so that

$$\sup_{a \in A} \operatorname{Re} f(a) = \alpha < f(x^{**}) \le \|f\|.$$

However

$$\alpha = \sup_{x \in b_1(X)} \operatorname{Re} f(x) = \|f\|.$$

This contradiction establishes the claim. ∎

3.3. Compactness in Weak Topologies

The unit ball of any infinite dimensional Banach space is not compact. A useful substitute is that the unit ball of a dual space is weak-$*$ compact. This sometimes allows us to use compactness in the Banach space setting.

3.3.1. BANACH-ALAOGLU THEOREM. *Let X be a Banach space. Then the closed unit ball of X^* is weak-$*$ compact.*

PROOF. For each $x \in X$, let $\mathbb{D}_x = \{z \in \mathbb{F} : |z| \le \|x\|\}$. These are compact sets, and so by Tychonoff's Theorem 1.7.7, $\mathcal{D} = \prod_{x \in X} \mathbb{D}_x$ is compact. Define

$$\Phi : (\overline{b_1(X^*)}, \tau_X) \to \mathcal{D} \quad \text{by} \quad \Phi(f)(x) = f(x) \text{ for } x \in X.$$

Claim 1: Φ is continuous. A basic open set in $\mathcal{D}$ is given by $x_1, \dots, x_n \in X$ and $U_i \subset \mathbb{D}_{x_i}$ open,

$$U = \{d \in \mathcal{D} : d(x_i) \in U_i,\ 1 \le i \le n\}.$$

Then

$$\Phi^{-1}(U) = \{f \in \overline{b_1(X^*)} : f(x_i) \in U_i\} = \overline{b_1(X^*)} \cap \bigcap_{i=1}^{n} f^{-1}(U_i).$$

This is the intersection of finitely many weak-$*$ open sets with $\overline{b_1(X^*)}$, so it is weak-$*$ open in the relative topology.

Claim 2: Φ is one to one. This is because X separates points of X^*.

Claim 3: $\Phi(\overline{b_1(X^*)})$ is closed in $\mathcal{D}$. Suppose that $f_\lambda \in \overline{b_1(X^*)}$ is a net such that $\Phi(f_\lambda) = d_\lambda$ converges in $\mathcal{D}$ to $d \in \mathcal{D}$. Then $d : X \to \mathbb{F}$ by evaluation at each coordinate, and $d(x) \in \mathbb{D}_x$, so $|d(x)| \leq \|x\|$. Check to see that d is linear: if $x, y \in X$ and $\alpha, \beta \in \mathbb{F}$, then

$$d(\alpha x + \beta y) = \lim_{\Lambda} f_\lambda(\alpha x + \beta y) = \lim_{\Lambda} \alpha f_\lambda(x) + \beta f_\lambda(y) = \alpha d(x) + \beta d(y).$$

Hence $d \in X^*$ and $\|d\| \leq 1$. Therefore $\Phi(\overline{b_1(X^*)})$ is a closed subset of a compact set, and thus it is compact.

Claim 4: Φ^{-1} is continuous on $\Phi(\overline{b_1(X^*)})$, so that Φ is a homeomorphism onto its image. A basic weak-$*$ open set in X^* has the form

$$V = \{f \in X^* : f(x_i) \in U_i, 1 \leq i \leq n\} \quad \text{for} \quad U_i \text{ open in } \mathbb{F}.$$

Observe that

$$\Phi(V \cap \overline{b_1(X^*)}) = \{d \in \mathcal{D} : d(x_i) \in U_i, 1 \leq i \leq n\} \cap \Phi(\overline{b_1(X^*)}).$$

This is relatively open in $\Phi(\overline{b_1(X^*)})$. Therefore $\overline{b_1(X^*)}$ is the continuous image of a compact set, and hence it is compact. ■

3.3.2. COROLLARY. *If X is a reflexive Banach space, then every norm-closed bounded convex set is weakly compact.*

PROOF. A closed convex set X is weakly closed by Proposition 3.2.14(1). Since it is bounded, it is contained in a large closed ball, which is weak-$*$ compact by the Banach-Alaoglu Theore. Since X is reflexive, the weak and weak-$*$ topologies are the same. So C is a weakly closed subset of a weakly compact set, and thus is weakly compact. ■

3.3.3. COROLLARY. $(\overline{b_1(X)}, \tau_{X^*})$ *is compact if and only if X is reflexive.*

PROOF. We just proved the only if direction, so suppose that $\overline{b_1(X)}$ is weakly compact. By Goldstine's Theorem, $\iota(\overline{b_1(X)}, \tau_{X^*})$ is dense in $(\overline{b_1(X^{**})}, \tau_{X^*})$. Since $\overline{b_1(X)}$ is weakly compact and ι is continuous, the image is compact and therefore closed. Hence $\overline{b_1(X^{**})} = \overline{b_1(X)}$, and $X^{**} = X$. ■

3.4. Extreme points

3.4.1. DEFINITION. Let K be a convex set in a vector space V. Then $F \subset K$ is a *face* of K if $x, y \in K$, $0 < t < 1$ and $tx + (1-t)y \in F$ implies that $x, y \in F$.

A point $x \in K$ is an *extreme point* of K if $\{x\}$ is a face. Let $\text{ext}(K)$ denote the set of all extreme points of K.

3.4.2. EXAMPLES.

(1) Let $K = \{(x,y) \in \mathbb{R}^2 : y \geq 0,\ x^2 + y^2 \leq 1\}$ be a semicircle. Then $F = [-1,1] \times \{0\}$ is a face. The extreme points are $\{(x,y) \in \mathbb{R}^2 : y \geq 0,\ x^2 + y^2 = 1\}$. Note that $\text{int}\, K$ is convex and has no extreme points. Also $\mathbb{R}$ is convex and has no extreme points. So we look for extreme points in closed bounded (often compact) convex sets.

(2) Let $C = \{(x,y,0) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$. Let $K = \text{conv}(C \cup \{(0,1,\pm 1)\})$. Then K is compact. $F = \{(0,1,z) : |z| \leq 1\}$ is a face. The extreme points are $\{(0,1,\pm 1)\} \cup C \setminus \{(0,1,0)\}$. Thus $\text{ext}\, K$ is not closed.

(3) A Banach space $(X, \|\cdot\|)$ is *strictly convex* if $\|x\| = \|y\| = \left\|\frac{x+y}{2}\right\|$ implies that $x = y$. Examples include Hilbert spaces and $L^p(\mu)$ for $1 < p < \infty$. If $1 = \|f\|_p = \|g\|_p$ and $2 = \|f+g\|_p \leq \|f\|_p + \|g\|_p = 2$, then equality in the triangle inequality shows that $f = g$. When X is strictly convex, $\text{ext}\, \overline{b_1(X)} = \{x \in X : \|x\| = 1\}$.

(4) Consider $K = \overline{b_1(L^1(0,1))}$. Let $f \in L^1(0,1)$ with $\|f\|_1 = 1$. Then there is some $t_0 \in (0,1)$ so that $\int_0^{t_0} |f|\, dx = \frac{1}{2}$. Let $f_1 = 2f\chi_{[0,t_0)}$ and $f_2 = 2f\chi_{[t_0,1)}$. Then $\|f_1\|_1 = \|f_2\|_1 = 1$ and $f = \frac{1}{2}(f_1 + f_2)$. Therefore K has no extreme points.

(4) Let $K \subset \mathbb{R}^n$ be a compact convex set. Then $\text{ext}\, K \neq \varnothing$ because $f(x) = \|x\|_2^2$ is continuous and takes its maximum on K at some point x_0. This is extreme in the unit ball of radius $\|x\|$, and so is also extreme in K.

It is a theorem of Carathéodory that if $E \subset \mathbb{R}^n$, every point in $\text{conv}(E)$ is a convex combination of at most $n+1$ points of E. This can be used to prove a theorem of Minkowski: in finite dimensions, a compact convex set $K = \text{conv}(\text{ext}\, K)$ without closure.

3.4.3. KREIN-MILMAN THEOREM. *Let X be a LCTVS, and let $K \subset X$ be a compact convex set. Then $\overline{\text{conv}}(\text{ext}\, K) = K$.*

PROOF. Let $\mathcal{F}$ be the set of closed faces of K ordered by containment; i.e. $F \leq G$ if $F \supset G$. This is a poset. Let's show that it is inductive. Suppose that $\mathcal{C} = \{F_\alpha : \alpha \in A\}$ is a chain of faces. Then $\mathcal{C}$ has the finite intersection property because the intersection of finitely many faces in $\mathcal{C}$ equals the smallest one. By compactness of K, $F = \bigcap_{\alpha \in A} F_\alpha$ is non-empty. Suppose that $x, y \in K$, $0 < t < 1$ and $tx + (1-t)y \in F$. Then $tx + (1-t)y \in F_\alpha$ for $\alpha \in A$. As each F_α is a face, $x, y \in F_\alpha$; and thus $x, y \in F$. Thus F is a face, and is an upper bound for $\mathcal{C}$. By Zorn's Lemma, $\mathcal{F}$ contains a maximal element, i.e. a minimal face, say F_0.

We claim that F_0 is a single point. It not, then since X^* separates points of X, there is a continuous linear functional f on X which is not constant on F_0. By scaling, we may suppose that Re f is not constant on F_0. Let $\alpha = \sup_{x \in F_0} \operatorname{Re} f(x)$ and $G = \{x \in F_0 : \operatorname{Re} f(x) = \alpha\}$. Then G is a proper face of F_0, contradicting minimality. Hence $F_0 = \{x_0\}$ is an extreme point. Therefore $\operatorname{ext} K \neq \varnothing$.

Let $L = \overline{\operatorname{conv}}(\operatorname{ext} K)$. This is a closed convex subset of K, and thus is compact. If $L \neq K$, pick $x \in K \setminus L$. By the Separation Theorem, there is an $f \in X^*$ so that

$$\sup_{x \in L} \operatorname{Re} f(x) = \alpha < \operatorname{Re} f(x) \leq \sup_{x \in K} \operatorname{Re} f(x) = \beta.$$

By compactness, the supremum is attained, so $F = \{x \in K : \operatorname{Re} f(x) = \beta\}$ is non-empty. It is easy to see that F is a compact face of K. Hence F has an extreme point x_0. A face of a face is a face, $x_0 \in \operatorname{ext} K$, contradicting the definition of L. Therefore $K = \overline{\operatorname{conv}}(\operatorname{ext} K)$. ■

3.4.4. COROLLARY (Krein-Milman)**.** *Let X be a Banach space. The closed unit ball of X^* is the weak-$*$ closed convex hull of its extreme points.*

PROOF. By the Banach-Alaoglu Theorem, $\overline{b_1(X^*)}$ is weak-$*$ compact and convex. Thus the Krein-Milman Theorem applies in (X^*, τ_X). ■

3.4.5. COROLLARY. *c_0 and $L^1(0,1)$ are not dual spaces.*

PROOF. The unit balls have no extreme points. ■

3.4.6. EXAMPLE. $C[0,1]^* = M[0,1]$ is the space of complex regular Borel measures on $[0,1]$. The point masses δ_x for $x \in [0,1]$ are extreme in $K = \overline{b_1(M[0,1])}$, as are $e^{i\theta}\delta_x$. If μ is a complex measure with $\|\mu\| = 1$ which is not supported on a single point, then there is a Borel set $A \subset [0,1]$ so that $0 < |\mu|(A) < 1$. Set $\mu_1(B) = \frac{1}{|\mu|(A)}\mu(B \cap A)$ and $\mu_2(B) = \frac{1}{|\mu|(A^c)}\mu(B \cap A^c)$. Then $\mu = |\mu|(A)\mu_1 + (1 - |\mu|(A))\mu_2$; so it is not extreme. Thus $\operatorname{ext} K = \{e^{i\theta}\delta_x : \theta \in \mathbb{R}, x \in [0,1]\}$. Now $\operatorname{conv}(\operatorname{ext} K)$ is the set of atomic measures of norm at most 1, not the whole ball. This is the intersection of K with the norm-closed subspace of atomic measures. The weak-$*$ closure is needed to obtain the other points in K.

Same works for $M(X) = C_0(X)^*$ where X is any locally compact Hausdorff space.

3.4.7. LEMMA. *If $K_1, \dots, K_n$ are compact convex sets in a LCTVS X, then* $\operatorname{conv}(K_1 \cup \dots \cup K_n)$ *is compact.*

PROOF. It is enough to establish the $n = 2$ case, since the general case then follows from repeated application.

Consider a point x in $\operatorname{conv}(K_1 \cup K_2)$ given by $x_j \in K_1 \cup K_2$, $0 < t_j < 1$ for $1 \le j \le m$, $\sum_{j=1}^m t_j = 1$ and $x = \sum_{j=1}^m t_j x_j$. Let $A_1 = \{j : x_j \in K_1\}$ and $A_2 = \{1, \dots, m\} \setminus A_1$. Let $s_1 = \sum_{j \in A_1} t_j$ and $s_2 = 1 - s_1 = \sum_{j \in A_2} t_j$. Then $y_i = \sum_{j \in A_i} \frac{t_j}{s_i} x_j$ is a convex combination of $\{x_j : j \in A_i\}$, and hence it belongs to K_i. Finally $x = s_1 y_1 + s_2 y_2$.

It follows from this analysis that the map $\Phi : K_1 \times K_2 \times [0,1] \to X$ given by $\Phi(x_1, x_2, s) = s x_1 + (1-s) x_2$ maps onto $\operatorname{conv}(K_1 \cup K_2)$. It is clear that Φ is a continuous map from a compact set into X. Therefore the image is compact. ■

There is a partial converse to the Krein-Milman Theorem.

3.4.8. THEOREM (Milman). *Let X be a LCTVS, and let $K \subset X$ be a compact convex set. Suppose that $E \subset K$ is closed and $\overline{\operatorname{conv}}(E) = K$. Then $\operatorname{ext} K \subset E$.*

PROOF. Suppose that $x_0 \in \operatorname{ext} K \setminus E$. Since E is closed, there is a neighbourhood $x_0 + U(F, r)$ which is disjoint from E. Now $\{x + U(F, r/2) : x \in E\}$ is an open cover of E, which is compact, so there is a finite subcover $x_i + U(F, r/2)$ for $1 \le i \le n$. Let $K_i = \overline{\operatorname{conv}}(E \cap x_i + U(F, r/2))$. This is contained in the convex set $\overline{x_i + U(F, r/2)}$, and thus $x_0 \notin K_i$. Then

$$\operatorname{conv}(E) \subset \operatorname{conv}(K_1 \cup \dots \cup K_n) \subset \overline{\operatorname{conv}}(E) = K.$$

By Lemma 3.4.7, the middle term is compact, so that $\operatorname{conv}(K_1 \cup \dots \cup K_n) = K$. In particular, x_0 is a convex combination of points in $K_1 \cup \dots \cup K_n$, which contradicts the fact that x_0 is extreme and not in this union. Therefore $\operatorname{ext} K \subset E$. ■

As an application of the new ideas in this course, we provide a proof of the Stone-Weierstrass Theorem due to de Bruijn.

3.4.9. STONE-WEIERSTRASS THEOREM. *Let X be a compact Hausdorff space. Suppose that A is a closed subalgebra of $C_{\mathbb{R}}(X)$ which separates points and doesn't vanish at any point in X. Then $A = C_{\mathbb{R}}(X)$.*

PROOF. Suppose that $A \ne C_{\mathbb{R}}(X)$. Then by the Hahn-Banach Theorem, $A^\perp$ is a non-zero weak-$*$ closed subspace of $C_{\mathbb{R}}(X)^*$. Let $K = \overline{b_1(C_{\mathbb{R}}(X)^*)} \cap A^\perp$. By the Banach-Alaoglu Theorem and the weak-$*$ closure of $A^\perp$, K is weak-$*$ compact and convex. By the Krein-Milman Theorem, K has an extreme point, say ψ. By the Riesz Representation Theorem, there is a finite regular signed Borel measure μ on X with $\|\mu\| = 1$ so that $\psi(f) = \int f \, d\mu$.

Claim: the support of μ is a single point. If not, let x, y be distinct points in $\operatorname{supp}(\mu)$. Since A separates points, there is some $f \in A$ so that $f(x) \ne f(y)$. Scaling and adding a constant produces a function $g \in A + \mathbb{R}\mathbf{1}$ so that $0 \le g \le 1$ and $g(x) \ne g(y)$. Then $g\mu$ is not a scalar multiple of μ. Note that $g\mu \in A^\perp$ since for $h \in A$, $\int hg \, d\mu = 0$ because $hg \in A$. Let $\mu_1 = g\mu / \|g\|_1$ and $\mu_2 =$

$(1-g)\mu/\|1-g\|_1$ where $\|g\|_1 = \int g\,d\mu$ and $\|g\|_1 + \|1-g\|_1 = 1$. Then μ_1 and μ_2 lie in $A^\perp$ and have norm 1, so belong to K. Then $\mu = \|g\|_1\mu_1 + \|1-g\|_1\mu_2$ is not extreme. This contradiction shows that μ is supported on a single point, say x_0. That is, $\mu = e^{i\theta}\delta_{x_0}$.

Since A does not vanish at x_0, there is some $f \in A$ so that $\psi(f) = \int f\,d\mu = e^{i\theta}f(x_0) \neq 0$. This contradicts the fact that $\psi \in A^\perp$. Therefore the original assumption was incorrect, and $A = C_{\mathbb{R}}(X)$. ∎

3.5. The Krein-Smulian Theorem

In this section, we establish a useful result for checking weak-$*$ closure of unbounded convex sets, including subspaces.

3.5.1. Definition. For $A \subset X$, define the *polar* of A to be

$$A^\circ = \{f \in X^* : \operatorname{Re} f(x) \leq 1 \text{ for } x \in A\}.$$

Note that A° is always a weak-$*$ closed convex set. In particular, observe that $\overline{b_r(X)}^\circ = \overline{b_{1/r}(X^*)}$.

3.5.2. Krein-Smulian Theorem. *Let X be a Banach space, and let $K \subset X^*$ be a convex set. Then K is weak-$*$ closed if and only of $K \cap \overline{b_r(X^*)}$ is weak-$*$ closed for all $r > 0$.*

Proof. The only if direction is trivial.

Suppose that $K \cap \overline{b_r(X^*)}$ is weak-$*$ compact for all $r > 0$. Then K is norm closed because norm convergent sequences are bounded, and so lie in the closed set $K \cap \overline{b_r(X^*)}$ for some large r. Let $f \in K^c$ and choose an $\varepsilon > 0$ so that $\overline{b_\varepsilon(f)} \cap K = \varnothing$. After translating and scaling, we may suppose that $f = 0$ and $\varepsilon = 1$, so $\overline{b_1(0)} \cap K = \varnothing$. We need to show that $0 \notin \overline{K}^{w*}$.

We will choose finite sets $\{A_n : n \geq 2\}$ in X so that

$$A_n \subset \overline{b_{\frac{1}{n-1}}(X)} \quad \text{and} \quad \overline{b_n(X^*)} \cap K \cap (A_2 \cup \cdots \cup A_n)^\circ = \varnothing.$$

We can begin with $A_1 = \varnothing$ and note that $\overline{b_1(X^*)} \cap K = \varnothing$. Suppose that $A_2, \ldots, A_n$ have been chosen with these properties. Let $D = \overline{b_{n+1}(X^*)} \cap K \cap (A_1 \cup \cdots \cup A_n)^\circ$. This is weak-$*$ compact and convex, and by assumption, $D \cap \overline{b_n(X^*)} = \varnothing$. Now $\overline{b_n(X^*)} = \overline{b_{1/n}(X)}^\circ$, so

$$\varnothing = D \cap \overline{b_n(X^*)} = D \cap \bigcap_{\substack{x \in X \\ \|x\| \leq 1/n}} \{x\}^\circ.$$

By compactness of D and the FIP, there is a finite subset $A_{n+1} \subset \overline{b_{1/n}(X)}$ so that $D \cap A_{n+1}^\circ = \varnothing$. This completes the induction.

List $\bigcup_{n\geq 2} A_n = \{x_1, x_2, \dots\}$ as a countable set with $\lim_{i\to\infty} x_i = 0$. Define a linear map $T : X^* \to c_0$ by $(Tf) = \big(f(x_i)\big)_{i\geq 1}$. We have $\|T\| = \sup_{i\geq 1} \|x_i\| \leq 1$.

If $f \in K$, then $\|f\| \leq n$ for n large enough, and thus $f \notin (A_1 \cup \cdots \cup A_n)^\circ$. Hence there is some $x_i \in A_1 \cup \cdots \cup A_n$ so that $\operatorname{Re} f(x_i) > 1$; so that $\|Tf\| > 1$. Therefore TK is convex and $TK \cap b_1(c_0) = \varnothing$. Since $b_1(c_0)$ is open, the Separation Theorem yields $\varphi = (a_1, a_2, \dots) \in c_0^* = l_1$ so that

$$\sup_{z\in b_1(c_0)} \operatorname{Re} \varphi(z) = \|\varphi\|_1 \leq \inf_{f\in K} \operatorname{Re} \varphi(Tf) = \inf_{f\in K} \operatorname{Re} \sum_{i\geq 1} a_i f(x_i) = \inf_{f\in K} \operatorname{Re} f(x_0)$$

where $x_0 = \sum_{i\geq 1} a_i x_i$, which converges since the $\{x_i, i \geq 1\}$ is bounded and $\sum |a_i| = \|\varphi\|_1 < \infty$. Therefore $x_0 \in X$ and $\operatorname{Re} f(x_0) \geq \|\varphi\|_1 > 0$ for all $f \in K$. It follows that $0 \notin \overline{K}^{w*}$. Hence K is weak-$*$ closed. ■

When K is a subspace, $K \cap \overline{b_r(X^*)} = r(K \cap \overline{b_1(X^*)})$. So the following is immediate.

3.5.3. COROLLARY. *If Y is a subspace of a dual space X^*, then Y is weak-$*$ closed if and only if $Y \cap \overline{b_1(X^*)}$ is weak-$*$ closed.*

3.5.4. COROLLARY. *For $\varphi \in X^{**}$, the following are equivalent:*

(1) $\varphi \in \iota X$.

(2) *φ is continuous on (X^*, τ_X); i.e. φ is weak-$*$ continuous.*

(3) *φ is weak-$*$ continuous on $\overline{b_1(X^*)}$.*

PROOF. The equivalence of (1) and (2) is Lemma 3.1.7. Clearly (2) implies (3). To show that (3) implies (2), Proposition 3.1.5 shows that it suffices to show that $\ker \varphi$ is weak-$*$ closed. Now (3) implies that $\ker \varphi \cap \overline{b_1(X^*)}$ is weak-$*$ closed. So the previous Corollary completes the proof. ■

3.5.5. DEFINITION. A subset $K \subset X$ of a LCTVS X is *sequentially closed* if whenever a sequence x_n in K has a limit x, then $x \in K$. So a subset K is *weak-$*$ sequentially closed* in a Banach dual space X^* if it is sequentially closed in (X^*, τ_X).

3.5.6. LEMMA. *Let X be a Banach space. Suppose that $Y \subset X^*$ is a separable subspace which separates points of X. Then $(\overline{b_1(X)}, \tau_Y)$ is metrizable.*

PROOF. Let $\{g_n : n \geq 1\}$ be a dense subset of $b_1(Y)$. Define a metric on $\overline{b_1(X)}$ by

$$d(x,y) = \sum_{n\geq 1} 2^{-n} |g_n(x-y)| \quad \text{for} \quad x, y \in \overline{b_1(X)}.$$

This is definite since Y separates points, and the triangle inequality follows since each term satisfies the triangle inequality. Let $\beta_r(x)$ denote the d-ball about x of radius r.

Suppose that $x \in \overline{b_1(X)}$, U is relatively open in $(\overline{b_1(X)}, \tau_Y)$ and $x \in U$. Find a finite $F = \{f_1, \dots, f_m\} \subset b_1(Y)$ and $r > 0$ so that $x + U(F, r) \cap \overline{b_1(X)} \subset U$. For each i, select n_i so that $\|f_i - g_{n_i}\| < \frac{r}{4}$ and let $N = \max\{n_i : 1 \leq i \leq m\}$. Suppose that $y \in \overline{b_1(X)}$ and $d(x,y) < \varepsilon = 2^{-1-N} r$. Then

$$\begin{aligned} |f_i(y) - f_i(x)| &\leq |(f_i - g_{n_i})(y)| + |f_{n_i}(y-x)| + |(f_i - g_{n_i})(x)| \\ &< \frac{r}{4} + 2^{n_i} d(y,x) + \frac{r}{4} \leq \frac{r}{2} + 2^N 2^{-1-N} r = r. \end{aligned}$$

Therefore $\beta_\varepsilon(x) \subset x + U(F, r) \subset U$.

Conversely, suppose that V is d-open and $x \in V$. Find an $\varepsilon > 0$ such that $\beta_\varepsilon(x) \subset V$. Find N so that $2^{1-N} < \frac{\varepsilon}{2}$, and let $F = \{g_1, \dots, g_N\}$. Suppose that $y \in x + U(F, \varepsilon/2)$. Then

$$\begin{aligned} d(y,x) &= \sum_{n\geq 1} 2^{-n} |f_n(y-x)| \\ &< \sum_{n=1}^{N} 2^{-n} \frac{\varepsilon}{2} + \sum_{n>N} 2^{-n} 2 \|f_n\| < \frac{\varepsilon}{2} + 2^{1-N} < \varepsilon. \end{aligned}$$

Thus $x + U(F, \varepsilon/2) \subset \beta_\varepsilon(x)$,

This shows that $(\overline{b_1(X)}, \tau_Y)$ and $(\overline{b_1(X)}, d)$ have the same open sets, and hence they determine the same topology. ■

3.5.7. COROLLARY. *Let X be a separable Banach space. Suppose that a convex $K \subset X^*$ is weak-$*$ sequentially closed. Then K is weak-$*$ closed.*

PROOF. By the Krein-Smulian Theorem, K is weak-$*$ closed if and only if $K \cap \overline{b_n(X^*)}$ is weak-$*$ closed for $n \geq 1$. But $(\overline{b_1(X^*)}, \tau_X)$ is metrizable by Lemma 3.5.6. In a metric space, a set is closed if and only if it is sequentially closed. ■

3.5.1. Preduals. A *predual* of a Banach space X is a Banach space Y such that Y^* is isometrically isomorphic to X. We have seen that some Banach spaces are not dual spaces. When a Banach space is a dual space, it may have a unique predual, or it may have many non-isomorphic preduals. Since every predual of X has a canonical imbedding into X^*, we can try to identify subspaces of X^* which are preduals of X.

3.5.8. EXAMPLE. Consider l_1. We know that $c_0^* = l_1$, but it has a plethora of preduals. We only mention a special class. Consider $C(K)$ where K is a countable locally compact metric space. Then the Riesz Representation Theorem states that $C(K)^* = M(K)$, the space of complex regular Borel measures on K. Since K is countable, every measure has the form $\mu = \sum_{x\in K} a_x \delta_x$ with $\|\mu\| = \sum_{x\in K} |a_x|$. It is clear that $M(K)$ is isometrically isomorphic to l_1. These spaces yield a variety of different Banach spaces. In particular the space

$$c = \{a_1, a_2, \dots) : \lim_{n\to\infty} a_n = a_0 \text{ exists}\} \simeq C(K) \quad \text{for} \quad K = \{0, \tfrac{1}{n} : n \geq 1\}.$$

is not isomorphic to c_0.

3.5.9. PROPOSITION. *Let X is a Banach space, then a closed subspace Y of X^* satisfies $Y^* = X$ if and only if*

(1) *Y norms X, and*

(2) *$\overline{b_1(X)}$ is τ_Y-compact.*

PROOF. If $X = Y^*$, then (1) is immediate and (2) follows from the Banach-Alaoglu Theorem.

Assume that (1) and (2) hold. Let $J : X \to Y^*$ by $Jx(y) = y(x)$. By (1), J is isometric: $\|Jx\| = \sup_{\|y\|\leq 1,\ y\in Y} |y(x)| = \|x\|$. Since $Y \subset X^*$, $(JY)_\perp = \{0\}$. Also J is continuous from (X, τ_Y) to (Y^*, τ_Y). Indeed, if E is a finite subset of Y, then $U(E, r) = \{f \in Y^* : |f(y)| < r,\ y \in E\}$ is a basis open neighbourhood of 0. Then $J^{-1}U(E, r) = \{x \in X : |y(x)| < r,\ y \in E\}$ is τ_Y-open in X. By (2), $\overline{b_1(X)}$ is τ_Y-compact. Therefore $J\overline{b_1(X)}$ is τ_Y-compact in Y^*; and hence it is weak-$*$ closed. The isometric property ensures that $J\overline{b_1(X)} = \overline{b_1(JX)} = JX \cap \overline{b_1(Y^*)}$. By the Krein-Smulian Theorem, JX is weak-$*$ closed in Y^*. Finally, by the Hahn-Banach Theorem, $JX = ((JX)_\perp)^\perp = \{0\}^\perp = Y^*$. So X is isometric to Y^*. ■

3.5.10. EXAMPLE. We will show that l_∞ has a unique predual, namely l_1. Let ε_n be the sequence in l_∞ with a 1 in the nth coordinate and zeros elsewhere; and let δ_n be the corresponding sequence in l_1. If $l_\infty = Y^*$, then $\overline{b_1(l_\infty)}$ is τ_Y-compact. Therefore $\overline{b_1(\varepsilon_n)} \cap \overline{b_1(-\varepsilon_n)} = \overline{b_1(\ker \delta_n)}$ is τ_Y-compact. By the Krein-Smulian Theorem, $\ker \delta_n$ is τ_Y- closed. By Proposition 3.1.5 and Lemma 3.1.7., $\delta_n \in Y$. Therefore $l_1 = \overline{\operatorname{span}\{\delta_n : n \geq 1\}} \subseteq Y$. If $Y \neq l_1$, there is some $0 \neq \varphi \in l_\infty = Y^*$ such that $\varphi|_{l_1} = 0$. This is impossible, and therefore $Y = l_1$.

Exercises for Chapter 3

1. Show that if $(x_\lambda)_{\lambda\in\Lambda}$ is a net in a Banach space X which converges weakly to x, then there is a sequence of points in the convex hull of the net which converges to x in norm.
HINT: show that the convex hull of the net intersects $b_{1/n}(x)$.

2. (a) Let X be a Banach space, and let Y be a closed subspace of X^* which norms X. Show that a *sequence* in X which converges in the τ_Y topology must be bounded.
(b) Let H be a Hilbert space. Show that if a *sequence* $T_n \in \mathcal{B}(H)$ converges to T in the weak operator topology, then $\{T_n\}$ is bounded. Note: the set of functionals determining this topology is not closed, so the proof from part (a) doesn't apply.

3. (a) Suppose that φ is a continuous functional on $(\mathcal{B}(H), \text{SOT})$. Show that there are vectors $x_i, y_i \in H$ for $1 \le i \le n$ so that $\varphi(T) = \sum_{i=1}^n \langle Tx_i, y_i\rangle$.
HINT: use that $\varphi^{-1}(\mathbb{D})$ is SOT-open to show that there are vectors x_i for $1 \le i \le n$ so that φ factors through the map $\Phi : \mathcal{B}(H) \to H^{(n)}$ given by $\Phi(T) = (Tx_1, Tx_2, \dots, Tx_n)$.
(b) Conclude that $(\mathcal{B}(H), \text{SOT})$ and $(\mathcal{B}(H), \text{WOT})$ have the same continuous linear functionals. Deduce that these two topologies have the same *closed convex* sets.
(c) Consider the operator M_z on $L^2(\mathbb{T})$ given by $M_z f = zf$, (where z is the function $z(e^{i\theta}) = e^{i\theta}$). Show that M_z^n converges to 0 in the WOT topology, but not in the SOT topology. So these two locally convex topologies are different.

4. Let X be a separable Banach space. Prove that $(\overline{b_1(X^*)}, \tau_X)$ is metrizable.
HINT: pick a countable dense subset $\{x_n : n \ge 1\}$ of $b_1(X)$. Define $d(\varphi, \psi) = \sum_{n\ge 1} 2^{-n} |(\varphi - \psi)(x_n)|$.

5. Let H be a Hilbert space. Prove that the closed unit ball of $\mathcal{B}(H)$ is compact in the weak operator topology. HINT: modify the proof of the Banach-Alaoglu Theorem.

6. Find all extreme points of the closed unit ball of (complex) l_p for $1 \le p < \infty$. Compute the norm-closed convex hull of ext$(\overline{b_1(l_p)})$.
HINT: equality in Minkowski's inequality.

7. Describe all extreme points of the closed unit ball of $C(K)$ where K is the Cantor set. What is the norm-closed convex hull of ext$(\overline{b_1(C(K))})$?

8. If $C_1, \dots, C_n$ are compact convex subsets of a LCTVS X, prove that conv$(C_1 \cup \dots \cup C_n)$ is compact.
HINT: write the convex hull as the continuous image of some compact set.

9. (a) Show that the space s of all complex sequences with the seminorms $p_n(x) = |x_n|$ for $n \geq 1$ is a Frechet space.
(b) Let s_0 be the subspace of s which are zero except finitely often. Put the τ_{s} topology on s_0. Show that $\mathrm{s}^* = \mathrm{s}_0$ and $\mathrm{s}_0^* = \mathrm{s}$.

10. The following deals with convex sets in finite dimensional spaces.
(a) Show that if $S \subset \mathbb{R}^n$, then every element of conv(S) is a convex combination of $n+1$ elements of S. HINT: if $s_1, \dots, s_m \in S$ with $m \geq n+2$, use the linear dependence of $s_i - s_m$ for $1 \leq i \leq m-1$.
(b) Hence show that the (non-closed) convex hull of a compact set is compact.
(c) Show that if $s_1, \dots, s_m$ are points in $\mathbb{R}^n$ with $m \geq n+2$, then you can split $\{1, \dots, m\}$ into two disjoint sets I and J so that conv$(I) \cap$ conv$(J) \neq \varnothing$.
(d) (Helly's Theorem) Let C_k be compact convex subsets of $\mathbb{R}^n$ for $k \geq 1$. Suppose that any $n+1$ of these sets have non-empty intersection. Show that $\cap_{k \geq 1} C_k$ is non-empty.
HINT: Use part (c) and induction to show that any m sets intersect.

11. (a) Show that a *sequence* $\mathbf{x}_n = (x_{n,i})_{i \geq 1}$ in l_p, $1 < p < \infty$, converges weakly to 0 if and only if (1) sup $\big\{ \|\mathbf{x}_n\|_p : n \geq 1 \big\} < \infty$ and (2) $\lim\limits_{n \to \infty} x_{n,i} = 0$ for all $i \geq 1$.
(b) Show that a *sequence* $\mathbf{x}_n = (x_{n,i})_{i \geq 1}$ in l_1 converges weakly if and only if it converges in norm.
(c) Find a *bounded net* in l_1 which converges weakly to 0 but does not converge in norm.

12. (a) A Banach space is *uniformly convex* if for every $\varepsilon > 0$, there is a $\delta > 0$ so that if $\|x\| \leq 1$, $\|y\| \leq 1$ and $\|x - y\| > \varepsilon$, then $\|(x+y)/2\| < 1 - \delta$. Show that Hilbert space is uniformly convex, but c_0 is not uniformly convex.
(b) Show that every uniformly convex Banach space is reflexive.
HINT: Given $z \in X^{**}$ with $\|z\| = 1$. Let $\varepsilon > 0$, get $\delta > 0$ from the uniformly convexity. Pick $f \in b_1(X^*)$ so that $|f(z) - 1| < \delta$. Let $C = \{x \in b_1(X) : |f(x) - 1| < \delta\}$. Show that z is in the weak* closure of C in X^{**}. Prove that $\|x - y\| < \varepsilon$ for all $x, y \in C$. Hence deduce that there is an $x \in C$ so that $\|z - x\| \leq \varepsilon$.

CHAPTER 4

Linear Operators

4.1. Adjoint Operators

4.1.1. DEFINITION. Let X, Y be Banach spaces. If $T \in \mathcal{B}(X, Y)$, define the *adjoint operator* $T^* \in \mathcal{B}(Y^*, X^*)$ by

$$(T^*g)(x) = g(Tx) \quad \text{for } g \in Y^* \text{ and } x \in X.$$

We sometimes write the pairing between a Banach space and its adjoint as a bilinear form: $\langle x, f \rangle = f(x)$ for $x \in X$ and $f \in X^*$. Unlike an inner product, this form is linear in each variable separately:

$$\langle a_1x_1 + a_2x_2, f \rangle = a_1\langle x_1, f \rangle + a_2\langle x_2, f \rangle \quad \text{for } x_1, x_2 \in X,\ a_1, a_2 \in \mathbb{F} \text{ and } f \in X^*$$

and

$$\langle x, b_1f_1 + b_2, f_2 \rangle = b_1\langle x, f_1 \rangle + b_2\langle x, f_2 \rangle \quad \text{for } x \in X,\ b_1, b_2 \in \mathbb{F} \text{ and } f_1, f_2 \in X^*.$$

Using this notation, we get

$$\langle x, T^*g \rangle = \langle Tx, g \rangle \quad \text{for } g \in Y^* \text{ and } x \in X.$$

4.1.2. THEOREM. *Let X, Y, Z be Banach spaces, and let $T \in \mathcal{B}(X, Y)$ and $S \in \mathcal{B}(Y, Z)$.*

(1) $\|T^*\| = \|T\|$.
(2) *The map $T \to T^*$ is linear.*
(3) $I_X^* = I_{X^*}$.
(4) $(ST)^* = T^*S^*$.
(5) *T^* is weak-$*$–weak-$*$ continuous; i.e. T^* is continuous from (Y^*, τ_Y) to (X^*, τ_X).*
(6) $T^{**}|_X = T$.

PROOF. (1) The Hahn-Banach Theorem is used for the second last equality:

$$\begin{aligned}\|T^*\| &= \sup_{\substack{\|g\|\le 1\\ g\in Y^*}} \|T^*g\| = \sup_{\substack{\|g\|\le 1\\ g\in Y^*}} \sup_{\substack{\|x\|\le 1\\ x\in X}} |\langle x, T^*g\rangle| \\ &= \sup_{\substack{\|x\|\le 1\\ x\in X}} \sup_{\substack{\|g\|\le 1\\ g\in Y^*}} |\langle Tx, g\rangle| = \sup_{\substack{\|x\|\le 1\\ x\in X}} \|Tx\| = \|T\|.\end{aligned}$$

(2) For $T_1, T_2 \in \mathcal{B}(X,Y)$, $a_1, a_2 \in \mathbb{F}$ and $g \in Y^*$,

$$\begin{aligned}\langle x, (a_1T_1 + a_2T_2)^*g\rangle &= \langle (a_1T_1 + a_2T_2)x, g\rangle = a_1\langle T_1x, g\rangle + a_2\langle T_2x_2, g\rangle \\ &= a_1\langle x, T_1^*g\rangle + a_2\langle x_2, T_2^*g\rangle = \langle x, a_1T_1^* + a_2T_2^*g\rangle.\end{aligned}$$

Therefore $(a_1T_1 + a_2T_2)^* = a_1T_1^* + a_2T_2^*$.

(3) $I_X^*f(x) = f(I_Xx) = f(x) = I_{X^*}f(x)$ for $x \in X$ and $f \in X^*$. Thus $I_X^* = I_{X^*}$.

(4) For $h \in Z^*$, $S^*h \in Y^*$ and $T^*S^*h \in X^*$. And for $x \in X$,

$$\langle x, T^*S^*h\rangle = \langle Tx, S^*h\rangle = \langle STx, h\rangle = \langle x, (ST)^*h\rangle.$$

Therefore $(ST)^* = T^*S^*$.

(5) Let $(g_\lambda)_\Lambda$ be a net in Y^* converging weak-$*$ to g. Then for $x \in X$,

$$\langle x, T^*g_\lambda\rangle = \langle Tx, g_\lambda\rangle \to \langle Tx, g\rangle = \langle x, T^*g\rangle.$$

Therefore $T^*g_\lambda \xrightarrow{w*} T^*g$. So T^* is weak-$*$–weak-$*$ continuous.

(6) For $x \in X$ and $g \in Y^*$,

$$\langle T^{**}x, g\rangle = \langle x, T^*g\rangle = \langle Tx, g\rangle.$$

Therefore $T^{**}|_X = T$. ■

4.1.3. EXAMPLE. Let $T : \mathbb{C}^m \to \mathbb{C}^n$. Fix a basis $e_1, \dots, e_m$ for $\mathbb{C}^m$ and $f_1, \dots, f_n$ for $\mathbb{C}^n$. Then T has an $n\times m$ matrix $T = \big[t_{ij}\big]$ so that $Te_j = \sum_{i=1}^n t_{ij}f_i$. Now $\mathbb{C}^{m*}$ has a *dual basis* $\delta_1, \dots, \delta_m$ given by $\delta_j(e_i) = \delta_{ij}$. Similarly $\mathbb{C}^{n*}$ has dual basis $\varepsilon_1, \dots, \varepsilon_n$. Then we calculate:

$$(T^*\varepsilon_j)(e_i) = \varepsilon_j(Te_i) = \varepsilon_j \sum_{k=1}^n t_{ki}f_k = t_{ji}.$$

Thus $T^*\varepsilon_j = \sum_{i=1}^m t_{ji}\delta_i$. Hence the matrix for T^* with respect to the dual bases is the $m \times n$ matrix $\big[t_{ji}\big] = \big[t_{ij}\big]^t$, which is the transpose of the matrix for T.

4.1.4. PROPOSITION. *Let X, Y be Banach spaces and let $T : X \to Y$ be a linear map. Then T is bounded (norm continuous) if and only if it is weak–weak continuous (i.e. from (X, τ_{X^*}) to (Y, τ_{Y^*})).*

PROOF. Suppose that T is norm continuous. Let $x_0 \in X$. If $G \subset Y^*$ is a finite set and $r > 0$, then $Tx_0 + U(G,r)$ is a basic τ_{Y^*}-open neighbourhood of Tx_0. Now

$$\begin{aligned} T^{-1}(Tx_0 + U(G,r)) &= x_0 + T^{-1}(\{y \in Y : |g(y)| < r,\ g \in G\}) \\ &= x_0 + \{x \in X : |g(Tx)| = (T^*g(x)| < r,\ g \in G\} \\ &= x_0 + U(T^*G, r) \end{aligned}$$

is τ_{X^*} open. Hence T is weak–weak continuous.

Conversely, suppose that T is weak–weak continuous. This is equivalent to saying that $g \circ T$ is weakly continuous on X for all $g \in Y^*$; i.e. $g \circ T \in X^*$. Therefore $\sup_{\|x\| \le 1} \|g(Tx)\| = |g \circ T| < \infty$ for all $g \in Y^*$. By the Banach-Steinhaus Theorem, $\|T\| = \sup_{\|x\|\le 1} \|Tx\| < \infty$. So T is bounded. ■

4.1.5. PROPOSITION. *Let X, Y be Banach spaces and let $T : Y^* \to X^*$ be a linear map. Then $T = S^*$ for some $S \in \mathcal{B}(X,Y)$ if and only if it is weak-∗–weak-∗ continuous (i.e. from (Y^*, τ_Y) to (X^*, τ_X)).*

PROOF. The direct direction is Theorem 4.1.2(5).

Suppose that T is weak-∗–weak-∗ continuous. Hence $\hat{x} \circ T$ is weak-∗ continuous on Y^* for every $x \in X$, so that $\hat{x} \circ T \in Y$. Therefore

$$\sup_{\|g\|\le 1} |(Tg)(x)| = \|\hat{x} \circ T\| < \infty \quad \text{for all} \quad x \in X.$$

By the Banach-Steinhaus Theorem, $\|T\| = \sup_{\|g\|\le 1} \|Tg\| < \infty$. So T is bounded.

Therefore $T^* : X^{**} \to Y^{**}$. We will show that $T^*X \subset Y$. Let $x \in X$ and $g \in Y^*$; and $T^*\hat{x} = y \in Y^{**}$. Then

$$y(g) = (T^*\hat{x})(g) = \hat{x}(Tg) = Tg(x).$$

If we show that y is a τ_Y-continuous functional on Y^*, then it belongs to Y. Let $(g_\lambda)_\Lambda$ be a net such which converges weak-∗ to g. Then since T is weak-∗–weak-∗ continuous, $(Tg_\lambda)_\Lambda$ converges weak-∗ to Tg. Thus

$$y(g_\lambda) = Tg_\lambda(x) \to Tg(x) = y(g).$$

This shows that y is a τ_Y continuous linear functional on Y^* and hence $y \in Y$; i.e. $T^*X \subset Y$.

Let $S = T^*|_X \in \mathcal{B}(X,Y)$. Then for $x \in X$ and $g \in Y^*$,

$$(S^*g)(x) = g(Sx) = g(T^*x) = (T^*\hat{x})(g) = \hat{x}(Tg) = (Tg)(x).$$

This is true for all $x \in X$ and $g \in Y^*$, and therefore $S^* = T$. ■

4.1.6. EXAMPLES.

(1) One might suspect that $\overline{\operatorname{Ran} T^*}$ is always weak-∗ closed, but this is not true. Let $T \in \mathcal{B}(l_1)$ given by $T(x_n) = (\frac{1}{n}x_n)$. Then it is easy to see that $T^* \in \mathcal{B}(l_\infty)$

also is given by $T^*(y_n) = (\frac{1}{n} y_n)$. Then $\operatorname{Ran} T^* \subset c_0$, and it is easy to check that $\overline{\operatorname{Ran} T^*} = c_0$. This is weak-$*$ dense in l_∞; and in particular is not weak-$*$ closed.

(2) Let X be a Banach space. Let $\iota_X : X \to X^{**}$ and $\iota_{X^*} : X^* \to X^{***}$ be the canonical injections. Then $\iota_X^* : X^{***} \to X^*$ is the restriction map to X, i.e., $\iota_X^* \Phi(x) = \Phi(\iota_X x) = \Phi|_X(x)$. Hence it it follows that $\iota_X^* \iota_{X^*} = I_{X^*}$. Let $P = \iota_{X^*} \iota_X^* \in \mathcal{B}(X^{***})$. Then

$$P^2 = \iota_{X^*}(\iota_X^* \iota_{X^*})\iota_X^* = \iota_{X^*} I_{X^*} \iota_X^* = P.$$

Moreover $\operatorname{Ran} P$ is contained in $\operatorname{Ran} \iota_{X^*}$ and if $\Phi = \iota_{X^*} f$, then

$$P\Phi = \iota_{X^*}(\iota_X^* \iota_{X^*}) f = \iota_{X^*} f = \Phi.$$

Thus P is the canonical projection of X^{***} onto X^*. Also $\|P\| \le \|\iota_{X^*}\| \, \|\iota_X^*\| = 1$, so that $\|P\| = 1$.

4.1.7. Lemma. *Let $T \in \mathcal{B}(X, Y)$. Then*

$$\ker T^* = (\operatorname{Ran} T)^\perp \quad \textit{and} \quad \ker T = (\operatorname{Ran} T^*)_\perp .$$

Proof. Let $g \in Y^*$. Then

$$g \perp \operatorname{Ran} T \iff 0 = \langle Tx, g\rangle = \langle x, T^* g\rangle \ \text{ for all } \ x \in X \iff T^* g = 0.$$

Similarly, for $x \in X$,

$$x \perp \operatorname{Ran} T^* \iff 0 = \langle x, T^* g\rangle = \langle Tx, g\rangle \ \text{ for all } \ g \in Y^* \iff Tx = 0.$$

The last equivalence is a consequence of the Hahn-Banach Theorem. ■

4.2. The Hilbert space Adjoint

When studying operators on a Hilbert space, there is a different operation which is also called the adjoint. This can lead to some confusion at first. The notion is based on Theorem 2.3.6 that the dual of a Hilbert space is conjugate linearly isomorphic to the original space.

4.2.1. Definition. Let H, K be Hilbert spaces. A *bounded sesquilinear form* on $H \times K$ is a map $[\cdot,\cdot] : H \times K \to \mathbb{F}$ which is linear in the first variable and conjugate linear in the second variable which satisfies $|[x,y]| \le C\|x\| \, \|y\|$ for all $x \in H$ and $y \in K$.

4.2.2. Proposition. *Let H, K be Hilbert spaces. Suppose that there is a sesquilinear form $[\cdot,\cdot]$ on $H \times K$ bounded by C. Then there is a unique operator B in $\mathcal{B}(K, H)$ so that $[x, y] = \langle x, By\rangle$ for $x \in H$ and $y \in K$. Moreover $\|B\| \le C$.*

PROOF. Fix $y \in K$, and define a linear functional $\varphi_y(x) = [x, y]$ on H. Note that

$$\|\varphi_y\| = \sup_{\|x\|\le 1} \big|[x,y]\big| \le \sup_{\|x\|\le 1} C\|x\|\,\|y\| = C\|y\|.$$

By Theorem 2.3.6, there is a unique vector $B(y) \in H$ so that $[x, y] = \langle x, B(y)\rangle$; and $\|B(y)\| = \|\varphi_y\| \le C\|y\|$. At this point, B is just a function from K to H. However it is clear that B is uniquely defined.

We verify that B is linear: if $y_1, y_2 \in K$ and $\lambda_1, \lambda_2 \in \mathbb{F}$, then

$$\begin{aligned}\langle x, B(\lambda_1 y_1 + \lambda_2 y_2)\rangle &= [x, \lambda_1 y_1 + \lambda_2 y_2] = \overline{\lambda_1}[x, y_1] + \overline{\lambda_2}[x, y_2] \\ &= \overline{\lambda_1}\langle x, B(y_1)\rangle + \overline{\lambda_2}\langle x, B(y_2)\rangle = \langle x, \lambda_1 B(y_1) + \lambda_2 B(y_2)\rangle.\end{aligned}$$

This holds for all $x \in H$, and thus $B(\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 B(y_1) + \lambda_2 B(y_2)$. That is, B is a linear map. We already have shown that $\|B\| \le C$. ∎

If $T \in \mathcal{B}(H, K)$, there is a sesquilinear form on $H \times K$ given by $[x, y] = \langle Tx, y\rangle$; and $[x, y]| \le \|T\|\,\|x\|\,\|y\|$. An immediate application of this proposition, there is a unique operator $B \in \mathcal{B}(K, H)$ so that $\langle Tx, y\rangle = \langle x, By\rangle$.

4.2.3. DEFINITION. Let H, K be Hilbert spaces. If $T \in \mathcal{B}(H, K)$, the *Hilbert space adjoint* or just *adjoint* of T is the unique operator $T^* \in \mathcal{B}(K, H)$ such that $\langle Tx, y\rangle = \langle x, T^* y\rangle$ for all $x \in H$ and $y \in K$.

The proof of the following proposition is very similar to the proof of Theorem 4.1.2. The details are left to the reader.

4.2.4. PROPOSITION. *Let H, K, L be Hilbert spaces. If $S, T \in \mathcal{B}(H, K)$ and $R \in \mathcal{B}(K, L)$, then*

(1) $\|T^*\| = \|T\|$.

(2) $(\lambda S + \mu T)^* = \bar{\lambda} S^* + \bar{\mu} T^*$ *for* $\lambda, \mu \in \mathbb{F}$. *So the adjoint is a conjugate linear map.*

(3) $I_H^* = I_H$.

(4) $(RT)^* = T^* R^*$.

(5) $T^{**} = T$.

4.2.5. EXAMPLE. Let $T \in \mathcal{B}(H)$ where $\dim H = n < \infty$. Fix an orthonormal basis $e_1, \dots, e_n$. Then T has a matrix $\big[t_{ij}\big]$ with respect to this basis. Then

$$\langle T^* e_j, e_i\rangle = \langle e_j, Te_i\rangle = \overline{t_{ji}} \quad \text{for} \quad 1 \le i, j \le n.$$

Therefore T^* has the matrix $\big[\,\overline{t_{ji}}\,\big]$, which is the conjugate transpose of $\big[t_{ij}\big]$.

4.3. Invertible Operators and the Spectrum

4.3.1. Definition. An operator $T \in \mathcal{B}(X, Y)$ is *bounded below* if there is some $\varepsilon > 0$ so that $\|Tx\| \geq \varepsilon \|x\|$ for $x \in X$.

4.3.2. Examples.
(1) Let $K \in \mathcal{B}(l_p)$ for $1 \leq p < \infty$ given by $K(x_n) = (\frac{1}{n} x_n)$. Then K is one-to-one, and the range of K is dense since it contains the standard basis $\{e_n : n \geq 1\}$. It is not bounded below since $\|Ke_n\| = \frac{1}{n} \to 0$. Also K is not surjective because $\operatorname{Ran} K$ does not contain $x = (n^{-1-1/p})$. You can easily check that $x \in l_p$ but $(n^{-1/p})$ does not. So K is not invertible.

(2) Let $S \in \mathcal{B}(l_p)$ for $1 \leq p < \infty$ given by $S(x_n) = (0, x_1, x_2, \dots)$, known as a *unilateral shift*. Then S is isometric: $\|Sx\|_p = \|x\|_p$. In particular, S is bounded below, and hence one-to-one. The range of S is $\{x = (x_n) : x_1 = 0\}$, which is a proper closed subspace. Thus S is not surjective, and hence is not invertible.

(3) Let $T \in \mathcal{B}(l_p)$ be given by $T(x_n) = (x_2, x_3, x_4, \dots)$. This is the *backward shift*. Then $\ker T = \mathbb{C}e_1$, and thus T is not invertible. However T is surjective. Moreover, $TS = I$. So S is left invertible and T is right invertible, but neither is invertible. Now $ST(x_1, x_2, x_3, \dots) = (0, x_2, x_3, \dots)$; so ST has kernel and proper closed range.

(4) Let $T \in \mathcal{M}_n = \mathcal{B}(\mathbb{C}^n)$, the space of $n \times n$ matrices. Then T is invertible if and only if it is one-to-one if and only if it is surjective. Thus if T is either left or right invertible, then T is invertible. The previous examples show that these results for invertible matrices all fail for operators on infinite dimensional spaces.

4.3.3. Proposition. *Let $T \in \mathcal{B}(X, Y)$. The following are equivalent:*

(1) *T is invertible.*
(2) *T is one-to-one and onto.*
(3) *T is bounded below and has dense range.*
(4) *T and T^* are bounded below.*
(5) *T^* is invertible.*

Proof. (1) implies (2) is clear; and (2) implies (1) is the Banach Isomorphism Theorem.

(1) implies (3) because $\operatorname{Ran} T = Y$ and $\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\| \, \|Tx\|$. Thus $\|Tx\| \geq \frac{1}{\|T^{-1}\|} \|x\|$ is bounded below.

(3) implies (2). Clearly bounded below implies one-to-one. We have that $\|Tx\| \geq \varepsilon_0\|x\|$ for some $\varepsilon_0 > 0$. Let $y \in Y$. Since T has dense range, there is a sequence $(x_n) \in X$ so that $Tx_n = y_n \to y$. Since y_n converges, it is a Cauchy sequence. So given $\varepsilon > 0$, there is an N so that $N \leq m < n$ implies that $\|T(x_m - x_n)\| = \|y_m - y_n\| < \varepsilon\varepsilon_0$. Therefore $\|x_m - x_n\| < \varepsilon$. So (x_n) is also Cauchy. Since X is complete, $x = \lim x_n$ exists. Therefore $Tx = \lim Tx_n = y$. Hence T is onto.

(1) implies (5). We have that $T^{-1}T = I_X$ and $TT^{-1} = I_Y$. Taking adjoints yields $T^*T^{-1*} = I_{X^*}$ and $T^{-1*}T^* = I_{Y^*}$. The first identity shows that T^* is surjective, and the second shows that it is injective. Thus T^* is invertible.

(5) implies (4). Since T^* is invertible, it is bounded below. Also T^{**} is invertible, and hence bounded below. Since $T = T^{**}|_X$, we see that T is bounded below.

(4) implies (3). We have T is bounded below; and $(\operatorname{Ran} T)^\perp = \ker T^* = \{0\}$ by Lemma 4.1.7 and the fact that T^* is bounded below. It follows from the Hahn-Banach Theorem that $\operatorname{Ran} T$ is dense because if $M = \overline{\operatorname{Ran} T}$ is a proper subspace, then there is a non-zero linear functional $g \in Y^*$ in $M^\perp$, and hence in $\ker T^*$. ■

When we discuss the spectrum of a linear operator, we will always work in complex vector spaces. In the study of matrices, the eigenvalues of a real matrix may be complex. Things are more convenient when we assume that $\mathbb{C}$ is our field.

4.3.4. Definition. Let $T \in \mathcal{B}(X)$ where X is a complex Banach space.

(1) The *spectrum* of T is $\sigma(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not invertible}\}$.

(2) The *resolvent* of T is $\rho(T) = \mathbb{C} \setminus \sigma(T)$.

(3) The *resolvent function* of T on $\rho(T)$ is $R(T, \lambda) = (\lambda I - T)^{-1}$.

(4) The *point spectrum* of T is $\sigma_p(T) = \{\lambda \in \mathbb{C} : \ker(\lambda I - T) \neq \{0\}\}$, the *eigenvalues* of T.

(5) The *approximate point spectrum* of T is

$$\pi(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not bounded below}\}.$$

(6) The *compression spectrum* of T is

$$\gamma(T) = \{\lambda \in \mathbb{C} : \operatorname{Ran}(\lambda I - T) \text{ is not dense in } X\}.$$

By Proposition 4.3.3, $\sigma(T) = \pi(T) \cup \gamma(T)$.

4.3.5. Proposition. *Let X be a Banach space. Then $\mathcal{B}(X)^{-1}$ is open and contains $b_1(I)$.*

PROOF. If $\|A\| < 1$, then $(I - A)^{-1} = \sum_{n\geq 0} A^n$. This series converges in $\mathcal{B}(X)$ because

$$\sum_{n\geq 0} \|A^n\| \leq \sum_{n\geq 0} \|A\|^n = \frac{1}{1 - \|A\|} < \infty.$$

Then one sees that

$$(I - A)\sum_{n\geq 0} A^n = \lim_{N\to\infty}(I - A)\sum_{n=0}^{N} A^n = \lim_{N\to\infty} I - A^{N+1} = I.$$

Similarly, $\sum_{n\geq 0} A^n(I - A) = I$. So $b_1(I) \subset \mathcal{B}(X)^{-1}$.

Now suppose that T is invertible and $\|A\| < 1/\|T^{-1}\|$. Then

$$T - A = T(I - T^{-1}A).$$

Since $\|T^{-1}A\| \leq \|T^{-1}\|\,\|A\| < 1$, it follows that $(I - T^{-1}A)$ is invertible. So $(T - A)^{-1} = (I - T^{-1}A)^{-1}T^{-1}$. Thus $b_{1/\|T^{-1}\|}(T) \subset \mathcal{B}(X)^{-1}$. Therefore $\mathcal{B}(X)^{-1}$ is open. ■

4.3.6. PROPOSITION. *Let $T \in \mathcal{B}(X)$. Then $\sigma(T)$ is a compact subset of $\overline{b_{\|T\|}(0)}$. Thus $\rho(T)$ is open.*

PROOF. Let $f : \mathbb{C} \to \mathcal{B}(X)$ by $f(\lambda) = \lambda I - T$. Clearly f is continuous, and thus $\rho(T) = f^{-1}(\mathcal{B}(X)^{-1})$ is open. Therefore $\sigma(T)$ is closed. If $|\lambda| > \|T\|$, then

$$\lambda I - T = \lambda(I - \lambda^{-1}T) \quad \text{and} \quad \|\lambda^{-1}T\| = \frac{\|T\|}{|\lambda|} < 1.$$

Therefore

$$(\lambda I - T)^{-1} = \lambda^{-1}\sum_{n\geq 0}(\lambda^{-1}T)^n = \sum_{n\geq 0}\lambda^{-n-1}T^n.$$

Thus $\sigma(T) \subset \overline{b_{\|T\|}(0)}$ and is closed; whence it is compact. ■

4.3.7. PROPOSITION. *The map $T \to T^{-1}$ is continuous on $\mathcal{B}(X)^{-1}$.*

PROOF. If $T_k \to I$, write $T_k = I - A_k$ where $\|A_k\| \to 0$. Once $\|A_k\| < 1$, we have that

$$\|I - (I - A_k)^{-1}\| = \Big\|\sum_{n\geq 1} A_k^n\Big\| \leq \sum_{n\geq 1}\|A_k\|^n = \frac{\|A_k\|}{1 - \|A_k\|} \to 0.$$

Hence the inverse map is continuous at I. Now suppose that T is invertible and $S_k \to T$. Then $S_kT^{-1} \to I$ since multiplication is continuous. Then $TS_k^{-1} \to I$ by continuity of the inverse at I. Left multiplying by T^{-1} yields $S_k^{-1} \to T^{-1}$. ■

4.3.8. EXAMPLES.

(1) Let $H = L^p(0,1)$ for $1 \le p < \infty$, and let $f \in L^\infty(0,1)$. Define the multiplication operator $M_f h = fh$. Recall that

$$\|f\|_\infty = \operatorname{ess\,sup}|f| = \sup\big\{r : m(\{x : |f(x)| > r\}) > 0\big\}.$$

Claim: $\|M_f\| = \|f\|_\infty$. Indeed, if $\|h\|_p \le 1$, then

$$\|M_f h\|_p = \int |f|^p |h|^p \, dm \le \|f\|_\infty^p \int |h|^p \, dm \le \|f\|_\infty^p.$$

So $\|M_f\| \le \|f\|_\infty$. Conversely, if $r < \|f\|_\infty$, then $A_r = \{x : |f(x)| > r\}$ has positive measure. Let $h = \dfrac{\overline{\operatorname{sign}(f)}}{m(A_r)^{1/p}} \chi_{A_r}$. Then $\|h\|_p = 1$ and

$$\|M_f h\|_p = \Big\| \frac{|f|}{\sqrt{m(A_r)}} \chi_{A_r} \Big\|_p > r\|h\|_p = r.$$

Therefore $\|M_f\| \ge \|f\|_\infty$, and hence we have equality.

Next note that if $g \in L^\infty$, then $M_f M_g = M_{fg} = M_g M_f$. So the map from $L^\infty(0,1)$ into $\mathcal{B}(L^p(0,1))$ is an isometric algebra isomorphism. Define

$$\operatorname{ess\,ran}(f) = \big\{\lambda \in \mathbb{C} : m\big(f^{-1}(b_\varepsilon(\lambda))\big) > 0 \text{ for all } \varepsilon > 0\big\}.$$

If $\lambda \notin \operatorname{ess\,ran}(f)$, then there is some $\varepsilon > 0$ so that $f^{-1}(b_\varepsilon(\lambda))$ has measure 0. Thus $g = (\lambda - f)^{-1}$ is essentially bounded by ε^{-1}, and thus lies in $L^\infty(0,1)$. Therefore $(\lambda I - M_f)M_g = M_g(\lambda I - M_f) = M_1 = I$. Consequently, $\sigma(M_f) \subset \operatorname{ess\,ran}(f)$. On the other hand, if $\lambda \in \operatorname{ess\,ran}(f)$, then $A_\varepsilon = f^{-1}(b_\varepsilon(\lambda))$ has positive measure for all $\varepsilon > 0$. Observe that

$$\|(\lambda I - M_f)\chi_{A_\varepsilon}\| = \|(\lambda - f)\chi_{A_\varepsilon}\| < \varepsilon\|\chi_{A_\varepsilon}\|.$$

Therefore $\lambda I - M_f$ is not bounded below, and hence it is not invertible. Hence $\sigma(M_f) = \operatorname{ess\,ran}(f)$. Note that the entire spectrum is approximate point spectrum.

The operator M_x has no point spectrum because if $xh = \lambda h$ a.e., we need $h = 0$ a.e.. Also $\operatorname{Ran}(\lambda I - M_x)$ is always dense since it is onto unless $\lambda \in [0,1]$, and in that case, $\operatorname{Ran}(\lambda I - M_x)$ contains $L^p(0, \lambda - \varepsilon) + L^p(\lambda + \varepsilon, 1)$ for all $\varepsilon > 0$. The union of these subspaces is dense in $L^p(0,1)$. So M_x has no compression spectrum. So $\sigma(M_x) = \pi(M_x) = [0,1]$ and $\sigma_p(M_x) = \varnothing = \gamma(M_x)$.

Consider $M_f^* \in \mathcal{B}(L^q(0,1)$ where $\frac{1}{p} + \frac{1}{q} = 1$. If $h \in L^p(0,1)$ and $k \in L^q(0,1)$, then

$$\langle M_f h, k\rangle = \int (fh)k \, dm = \int h(fk)\, dm = \langle h, M_f^* k\rangle.$$

Thus $M_f^* = M_f \in \mathcal{B}(L^q(0,1)$.

In the case of $p = 2$, we are in the Hilbert space situation where the adjoint is defined differently. For $h, k \in L^2(0,1)$

$$\langle M_f h, k\rangle = \int (fh)\overline{k}\, dm = \int h(f\overline{k})\, dm = \langle h, \overline{f}k\rangle = \langle h, M_{\overline{f}} k\rangle.$$

Thus $M_f^* = M_{\bar f}$ in this situation. Since $M_f M_f^* = M_{|f|^2} = M_f^* M_f$, we say that M_f is a *normal* operator.

(2) A similar but easier analysis can be carried out for the diagonal operators on l_p for $1 \le p < \infty$. Given a bounded sequence $(d_n) \in l_\infty$, define $D \in \mathcal{B}(l_p)$ by $D(a_n) = (d_n a_n)$. It is left to the reader to show that $\sigma_p(D) = \{d_n : n \ge 1\}$ and $\sigma(D) = \overline{\{d_n : n \ge 1\}}$. We will write $D = \operatorname{diag}(d_1, d_2, d_3, \dots)$ since the 'matrix' of D is diagonal.

(3) Shifts. Let S and T be the forward and backward shifts on l_p for $1 \le p < \infty$ defined in Example 4.3.2(2,3). We saw that S is an isometry with proper closed range, and so is not invertible. Also $TS = I$ but $\ker ST = \mathbb{C}e_1$. We claim that $\sigma(S) = \sigma(T) = \overline{\mathbb{D}}$. Since $\|S\| = \|T\| = 1$, their spectra are contained in the closed disc. If $|\lambda| < 1$, let $x_\lambda = (1, \lambda, \lambda^2, \dots)$. Then

$$\|x_\lambda\|_p^p = \sum_{n\ge 0} |\lambda|^{np} = \frac{1}{1-|\lambda|^p} < \infty.$$

Observe that $Tx_\lambda = (\lambda, \lambda^2, \lambda^3, \dots) = \lambda x_\lambda$. Hence $\sigma_p(T) \supset \mathbb{D}$, and hence $\sigma(T) = \overline{\mathbb{D}}$. Even though l_p is separable, T has uncountably many eigenvalues. However S_p does not have any eigenvalues at all: if $\lambda(a_n) = S(a_n) = (0, a_1, a_2, \dots)$, then $\lambda = 0$; but S is an isometry, so has no kernel.

Next we identify the adjoint. We will write S_p and T_p for the shifts on l_p. We calculate the adjoints of S_p and T_p. Suppose that $h = (a_n) \in l_p$ and $k = (b_n) \in l_q$, where where $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\langle S_p h, k\rangle = \sum_{n=2}^{\infty} a_n b_{n+1} = \langle (a_n), (b_2, b_3, \dots)\rangle = \langle h, T_q k\rangle.$$

Therefore $S_p^* = T_q$. The calculation to show that $T_p^* = S_q$ is similar. Now compute for $\lambda \in \mathbb{D}$ and $h \in l_p$,

$$\langle (\lambda I - S_p)h, x_\lambda\rangle = \langle h, (I - T_q)x_\lambda\rangle = 0.$$

(You can check that $T_\infty x_\lambda = \lambda x_\lambda$ in the $p = 1$ case.) Therefore $\operatorname{Ran}(\lambda I - S_p)$ annihilates the non-zero vector $x_\lambda \in l_q$, and therefore the closure of this range is proper. Hence $\lambda I - S_p$ is not invertible. So $\gamma(S_p) \supset \mathbb{D}$ and $\sigma(S_p) = \overline{\mathbb{D}}$.

4.4. Analyticity and the Resolvent

The theory of analytic functions plays an important role in functional analysis. There are at least two natural definitions of what that should mean.

4.4.1. DEFINITION. Let $\Omega \subset \mathbb{C}$ be open., and let X be a Banach space. A function $f : \Omega \to X$ is *weakly analytic* if $\varphi \circ f : \Omega \to \mathbb{C}$ is analytic for all $\varphi \in X^*$. The function f is *strongly analytic* if for every $z_0 \in \Omega$, there is a sequence $(x_n)_{n\geq 0} \subset X$ so that $f(z) = \sum_{n\geq 0} x_n(z-z_0)^n$ on a ball $b_r(z_0)$ for $r > 0$.

Clearly a strongly analytic function is weakly analytic. The converse is also true. We don't need this for our study of the spectrum because we can prove directly that the resolvent function is strongly analytic.

4.4.2. THEOREM. *Let X be a Banach space, and let $T \in \mathcal{B}(X)$. Then*

(1) $\dfrac{R(T,\lambda) - R(T,\mu)}{\lambda - \mu} = -R(T,\lambda)R(T,\mu)$ *for* $\lambda, \mu \in \rho(T)$.

(2) $\lambda \to R(T,\lambda)$ *is strongly analytic.*

(3) $\lim_{|\lambda|\to\infty} R(T,\lambda) = 0$.

PROOF. Note that $\lambda I - T$ and $(\mu I - T)^{-1}$ commute for $\lambda \in \mathbb{C}$ and $\mu \in \rho(T)$. Thus (1) follows from rearranging the identity

$$(R(T,\lambda) - R(T,\mu))(\lambda I - T)(\mu I - T) = (\mu I - T) - (\lambda I - T) = \mu - \lambda.$$

For (2), for $z_0 \in \rho(T)$ and $|w| < \|R(T,z_0)\|^{-1}$,

$$\begin{aligned}\big((z_0+w)I - T\big)^{-1} &= \big((z_0 I - T) - wI\big)^{-1} = \big((z_0 I - T)(I - wR(T,z_0))\big)^{-1} \\ &= R(T,z_0)\big(I - wR(T,z_0)\big)^{-1} \\ &= R(T,z_0)\sum_{n\geq 0} R(T,z_0)^n w^n = \sum_{n\geq 0} R(T,z_0)^{n+1} w^n.\end{aligned}$$

(3) For $|\lambda| > \|T\|$, we have

$$\|R(T,\lambda)\| = \Big\|\sum_{n\geq 0} \lambda^{-n-1} T^n\Big\| \leq \frac{1}{|\lambda| - \|T\|}.$$

This converges to 0 as $|\lambda| \to \infty$. ■

4.4.3. THEOREM. *Let X be a Banach space. If $T \in \mathcal{B}(X)$, then $\sigma(T)$ is non-empty.*

PROOF. If $\sigma(T)$ is empty, then $R(T,\lambda)$ is an entire function (an analytic function defined on all of $\mathbb{C}$). Since $\lim_{|\lambda|\to\infty} \|R(T,\lambda)\| = 0$, we see that $R(T,\lambda)$ is bounded. Therefore for each $\varphi \in X^*$, $\varphi \circ R(T,\lambda)$ is a bounded entire function. Thus this function is constant by Liouville's Theorem. By the Hahn-Banach Theorem, $R(T,\lambda)$ is constant. This is absurd, and so the spectrum is non-empty. ■

4.4.4. PROPOSITION. *Let $T \in \mathcal{B}(X)$. If $\lambda \in \rho(T)$, then*

$$\|R(T,\lambda)\| \geq \frac{1}{\operatorname{dist}(\lambda, \sigma(T))}.$$

Hence the boundary of the spectrum, $\partial\sigma(T)$, is in the approximate point spectrum $\pi(T)$.

PROOF. Pick $\lambda_0 \in \sigma(T)$ so that $|\lambda - \lambda_0| = \operatorname{dist}(\lambda, \sigma(T))$. Then

$$(\lambda_0 I - T)(\lambda I - T)^{-1} = (\lambda_0 - \lambda)(\lambda I - T)^{-1} + I$$

is not invertible. It follows that $|\lambda_0 - \lambda| \, \|\lambda I - T)^{-1}\| \geq 1$. Therefore

$$\|\lambda I - T)^{-1}\| \geq \frac{1}{|\lambda_0 - \lambda|} = \frac{1}{\operatorname{dist}(\lambda, \sigma(T))}.$$

Fix $\lambda_0 \in \partial\sigma(T)$. For $\varepsilon > 0$ there is a point $\lambda \in \rho(T)$ with $|\lambda - \lambda_0| < \varepsilon$. Hence $\|R(T,\lambda)\| > 1/\varepsilon$. So there is some vector $x \in X$ with $\|x\| = 1$ and $\|y\| = \|R(T,\lambda)X\| > 1/\varepsilon$. Hence

$$\frac{\|(\lambda I - T)y\|}{\|y\|} = \frac{\|x\|}{\|y\|} < \varepsilon.$$

Therefore

$$\frac{\|(\lambda_0 I - T)y\|}{\|y\|} \leq \frac{|\lambda_0 - \lambda|\|y\| + \|x\|}{\|y\|} < |\lambda_0 - \lambda| + \varepsilon < 2\varepsilon.$$

Hence $\lambda_0 I - T$ is not bounded below. ■

4.4.5. THEOREM (Spectral Mapping for Rational Funcions). *Let $f(z) = \frac{p(z)}{q(z)}$ be a rational function with no poles in $\sigma(T)$. Then $\sigma(f(T)) = f(\sigma(T))$.*

PROOF. We may suppose that p and q have no common factors. Write $q(z) = \prod_{i=1}^{n}(z - \alpha_i)$ where $\alpha_i \in \rho(T)$. Therefore $q(T)$ is invertible, so that $f(T) = p(T)q(T)^{-1}$ is defined. For $\lambda \in \mathbb{C}$,

$$\lambda - f(z) = \frac{p_\lambda(z)}{q(z)} = \frac{(z - \beta_1)\dots(z - \beta_m)}{(z - \alpha_1)\dots(z - \alpha_n)}.$$

Thus $\lambda I - T = \prod_{j=1}^{m}(T - \beta_j I)\ q(T)^{-1}$. Therefore $\lambda \in \sigma(f(T))$ if and only if some $\beta_j \in \sigma(T)$ if and only if there is $\beta \in \sigma(T)$ so that $0 = \dfrac{p_\lambda(\beta)}{q(\beta)} = \lambda - f(\beta)$ if and only if $\lambda \in f(\sigma(T))$. ■

4.4.6. DEFINITION. The *spectral radius* of T is

$$\operatorname{spr}(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

4.4.7. THEOREM. *Let $T \in \mathcal{B}(X)$. Then* $\operatorname{spr}(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$.

PROOF. By the Spectral Mapping Theorem, $\sigma(T^n) = \sigma(T)^n$, and hence $\operatorname{spr}(T^n) = \operatorname{spr}(T)^n$. Therefore

$$\operatorname{spr}(T) = \operatorname{spr}(T^n)^{1/n} \le \|T^n\|^{1/n}.$$

Hence $\operatorname{spr}(T) \le \inf_{n\ge1} \|T^n\|^{1/n}$.

Now $R(T,\lambda)$ is analytic on $\rho(T) \supset \{\lambda : |\lambda| > \operatorname{spr}(T)\}$. The Laurent expansion about ∞ is the series $R(T,\lambda) = \sum_{n\ge0} T^n\lambda^{-n-1}$. We know that this is valid for $|\lambda| > \|T\|$. However the function is analytic on the larger annulus $|\lambda| > \operatorname{spr}(T)$. So by analogy with the scalar case, we expect that this should converge absolutely and uniformly if $|\lambda| \ge r > \operatorname{spr}(T)$.

To see that this is indeed correct, note that for $\varphi \in X^*$,

$$\varphi(R(T,\lambda)) = \sum_{n\ge0} \varphi(T^n)\lambda^{-n-1}$$

is analytic on $\{\lambda : |\lambda| > \operatorname{spr}(T)\}$, and therefore this series converges absolutely for $|\lambda| > \operatorname{spr}(T)$. In particular, $|\varphi(T^n)\lambda^{-n-1}| \to 0$; whence $|\varphi(T^n)\lambda^{-n-1}| \le C$ for $n \ge 0$. Therefore $|\varphi(T^n)| \le Cr^{n+1}$ for any $r > \operatorname{spr}(T)$. Rearranging, we get that $\sup_{n\ge0} \left|\varphi\left(\frac{T^n}{r^n}\right)\right| \le Cr < \infty$. By the Banach-Steinhaus Theorem, we get $\sup_{n\ge0} \left\|\frac{T^n}{r^n}\right\| = C' < \infty$. Therefore $\|T^n\| \le C'r^n$; whence $\limsup \|T^n\|^{1/n} \le r$ for any $r > \operatorname{spr}(T)$. This means that

$$\limsup \|T^n\|^{1/n} \le \operatorname{spr}(T) \le \liminf \|T^n\|^{1/n}.$$

Therefore $\operatorname{spr}(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$. ■

We prove a lemma which will prove useful later.

4.4.8. LEMMA. *Suppose that $X = Y \dotplus Z$, and $T \in \mathcal{B}(X)$ satisfies $TY \subset Y$ and $TZ \subset Z$. Then $\sigma(T) = \sigma(T|_Y) \cup \sigma(T|_Z)$.*

PROOF. If $\lambda \in \rho(T)$, then $(\lambda I - T)Y \subset Y$ and $(\lambda I - T)Z \subset Z$. But $\lambda I - T$ is surjective, and hence $(\lambda I - T)Y = Y$ and $(\lambda I - T)Z = Z$. Also $\ker(\lambda I - T) = \{0\}$, so that $(\lambda I - T)|_Y$ is a continuous bijection on Y and $(\lambda I - T)|_Z$ is a continuous bijection on Z. So both are invertible by Banach's Isomorphism Theorem. Hence $\rho(T) \subset \rho(T|_Y) \cap \rho(T|_Z)$.

Conversely, suppose that $\lambda \in \rho(T|_Y) \cap \rho(T|_Z)$. Let $A = (\lambda I - T)|_Y^{-1} \in \mathcal{B}(Y)$ and $B = (\lambda I - T)|_Z^{-1} \in \mathcal{B}(Z)$. Define $S \in \mathcal{B}(X)$ by $S(y+z) = Ay + Bz$ for $x = y + z \in X$. Then for $y \in Y$ and $z \in Z$,

$$S(\lambda I - T)(y+z) = A(\lambda I - T)|_Y y + B(\lambda I - T)|_Z z = y + z$$

and

$$(\lambda I - T)S(y+z) = (\lambda I - T)|_Y Ay + (\lambda I - T)|_Z Bz = y + z.$$

Thus $S = (\lambda I - T)^{-1}$. Therefore $\rho(T|_Y) \cap \rho(T|_Z) \subset \rho(T)$. ■

Exercises for Chapter 4

1. If $A, B, C, D \in \mathcal{B}(\mathcal{H})$, then $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is a bounded operator on the Hilbert space $\mathcal{H} \oplus \mathcal{H}$ given by $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Ax + By \\ Cx + Dy \end{bmatrix}$. Likewise let $S = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$.

(a) Show that every operator T in $\mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ has this form, and show that TS is computed by matrix multiplication.
(b) Find T^*. (This is the Hilbert space adjoint.)
(c) Find all operators $E \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ such that $E = E^2$ and $\operatorname{Ran} E = \mathcal{H} \oplus \{0\}$.
(d) If A is invertible, show that T is invertible if and only if $D - CA^{-1}B$ is also invertible.
HINT: find X and Y to factor $T = \begin{bmatrix} A & 0 \\ C & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$.

2. Prove that a weakly analytic function is strongly analytic as follows.
(a) For $z_0 \in \overline{b_r(z_0)} \subset \Omega$, show that $\left\{\varphi\left(\frac{f(z_0+w)-f(z_0)}{w}\right) : 0 < |w| \le r\right\}$ is a bounded set for each $\varphi \in X^*$.

(b) Show that f is continuous on $\overline{b_r(z_0)}$.

(c) Define $x_n = \frac{1}{2\pi r^n} \int_0^{2\pi} f(z_0 + re^{i\theta}) e^{-in\theta}\, d\theta$ as a Riemann integral.

(d) Show that $f(z_0 + w) = \sum_{n \ge 0} x_n w^n$ for $|w| < r$.

3. Recall that $T \in \mathcal{B}(l_1)$ has a matrix $\left[t_{ij}\right]$ such that the columns are uniformly bounded Chapter 2, Exercise (2)). Show that T is the adjoint of an operator $S \in B(c_0)$ if and only if the rows of the matrix are in c_0.

4. (a) A weighted shift on a Hilbert space H acts on an orthonormal basis $\{e_n : n \ge 1\}$ by $We_n = w_n e_{n+1}$ for $n \ge 1$, where $\{w_n\}$ is a bounded sequence in $\mathbb{C}$. Compute $\|W\|$.
(b) Show that there is a diagonal unitary U_θ so that $U_\theta W U_\theta^* = e^{i\theta} W$ for each $\theta \in (0, 2\pi)$. Deduce that $\sigma(W)$ has circular symmetry.
(c) The Kakutani shift W_K has weights $w_n = 1/\gcd(n, 2^n)$ for $n \ge 1$. Compute $\operatorname{spr}(W_K)$.
(d) Let W_k be the weighted shift with weights $w_n = 1/\gcd(n, 2^n)$ if 2^k doesn't divide n, and $w_{2^k m} = 0$ for $m \ge 1$. Compute the spectral radius of W_k.
(e) Show that $\lim_{k \to \infty} W_k = W_K$. Deduce that spectral radius is not continuous on $\mathcal{B}(H)$.

.

CHAPTER 5

Compact Operators

In this chapter, X, Y, Z will denote Banach spaces; and H, K will be Hilbert spaces.

5.1. Compact Operators

5.1.1. DEFINITION. Let X, Y be Banach spaces. An operator $T \in \mathcal{B}(X,Y)$ is a *compact operator* if $\overline{Tb_1(X)}$ is compact in Y. We write $\mathcal{K}(X,Y)$ for the set of all compact operators in $\mathcal{B}(X,Y)$, and we write $\mathcal{K}(X)$ for $\mathcal{K}(X,X)$.

5.1.2. THEOREM. *$\mathcal{K}(X,Y)$ is a norm closed $\mathcal{B}(Y)$–$\mathcal{B}(X)$ bimodule. In particular, $\mathcal{K}(X)$ is a closed 2-sided ideal of $\mathcal{B}(X)$.*

PROOF. Let $K_1, K_2 \in \mathcal{K}(X,Y)$ be compact operators, and let $C_i = \overline{K_i b_1(X)}$. Then $C_1 \times C_2$ is compact. If $\lambda_1, \lambda_2 \in \mathbb{C}$, then $\lambda_1 C_1 + \lambda_2 C_2$ is compact because it is the continuous image of $C_1 \times C_2$ under the map $(x,y) \to \lambda_1 x + \lambda_2 y$. Therefore $\overline{(\lambda_1 K_1 + \lambda_2 K_2) b_1(X)} \subset \lambda_1 C_1 + \lambda_2 C_2$. A closed subset of a compact set is compact, and therefore $\lambda_1 K_1 + \lambda_2 K_2$ is compact. Hence $\mathcal{K}(X,Y)$ is a subspace of $\mathcal{B}(X,Y)$.

If $S \in \mathcal{B}(Y)$, $T \in \mathcal{B}(X)$ and $K \in \mathcal{K}(X,Y)$, then

$$\overline{SKTb_1(X)} \subset \overline{SKb_{\|T\|}(X)} = S\|T\|\overline{Kb_1(X)}$$

is the continuous image of a compact set, and hence is compact. Thus $\mathcal{K}(X,Y)$ is a $\mathcal{B}(Y)$–$\mathcal{B}(X)$ bimodule.

Finally we show that $\mathcal{K}(X,Y)$ is norm closed. So suppose that $K_n \in \mathcal{K}(X,Y)$ and $K_n \to K \in \mathcal{B}(X,Y)$. Let $\varepsilon > 0$. Pick N so that $\|K - K_N\| < \varepsilon/3$. Since $K_N b_1(X)$ has compact closure, it is totally bounded. Thus we can pick $y_i = K_N x_i$ with $\|x_i\| < 1$ for $1 \le i \le m$ to be a finite $\varepsilon/3$-net for $K_N b_1(X)$. Let $y_i' = Kx_i$ in $Kb_1(X)$. For any $x \in b_1(X)$, there is some i so that $\|K_N x - K_N x_i\| < \varepsilon/3$. Hence

$$\begin{aligned} \|Kx - y_i'\| &\le \|(K - K_N)x\| + \|K_N x - K_N x_i\| + \|(K_N - K)x_i\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus y_i' for $1 \le i \le m$ is an ε-net for $Kb_1(X)$; and therefore $\overline{Kb_1(X)}$ is compact. Hence K is a compact operator. ■

5.1.3. EXAMPLE. Finite rank operators are compact. Suppose $F \in \mathcal{B}(X,Y)$ has finite rank. Then $\operatorname{Ran} F = M$ is an n-dimensional subspace of Y, and therefore it is closed. Hence $\overline{Fb_1(X)}$ is a closed bounded subset of $M \simeq \mathbb{C}^n$. Thus it is compact. So F is compact.

We can deduce that any limit of finite rank operators is also compact. Consider a diagonal operator $D = \operatorname{diag}(d_1, d_2, \dots)$ on l_p for $1 \le p \le \infty$ is given by $D(a_n) = (d_n a_n)$. We have seen that D is bounded if and only if $\sup |d_n| < \infty$. We claim that it is compact if and only of $\lim_{n\to\infty} d_n = 0$. If the limit exists and is 0, then given $\varepsilon > 0$, there is an N so that $|d_n| < \varepsilon$ for all $n > N$. The diagonal operator $D_N = \operatorname{diag}(d_1, \dots, d_N, 0, 0 \dots)$ is finite rank, and thus compact. Also $\|D - D_N\| = \sup_{n>N} |d_n| < \varepsilon$. Hence $D = \lim_{n\to\infty} D_n$ is compact.

On the other hand, if 0 is not the limit, then there is some $\delta > 0$ and a sequence $n_i \to \infty$ so that $|d_{n_i}| \ge \delta$ for $i \ge 1$. The vectors $De_{n_i} = d_{n_i} e_{n_i}$ all belong to $\overline{Db_1(X)}$ and $\|De_{n_i} - De_{n_j}\| = \|d_{n_i} e_{n_i} - d_{n_j} e_{n_j}\| \ge 2^{1/p}\delta$. Therefore $\overline{Db_1(X)}$ is not totally bounded, because there is no finite δ net. So D is not compact.

5.1.4. EXAMPLE. Integral Operators. Let $k(x,y) \in L^2((0,1)^2)$, and define an operator $K \in \mathcal{B}(L^2(0,1))$ by

$$Kh(x) = \int_0^1 h(y) k(x,y)\, dy.$$

This is a *Hilbert-Schmidt integral operator*. Compute using Cauchy-Schwarz

$$\begin{aligned}\|Kh\|_2^2 &= \int_0^1 \Big| \int_0^1 h(y)k(x,y)\,dy \Big|^2 dx \le \int_0^1 \big(\|h\|_2 \, \|k(x,\cdot)\|_2\big)^2 dx \\ &= \|h\|_2^2 \int_0^1 \int_0^1 |k(x,y)|^2 \,dy\,dx = \|k\|_2^2 \, \|h\|_2^2.\end{aligned}$$

Therefore $\|K\| \le \|k\|_2$.

Let $\{e_i(x) : i \ge 1\}$ be an orthonormal basis for $L^2(0,1)$. Then $\{e_i(x)e_j(y) : i,j \ge 1\}$ is an orthonormal basis for $L^2((0,1)^2)$. Write $k = \sum_{i,j=1}^\infty a_{ij} e_i(x) e_j(y)$, where $\|k\|_2^2 = \sum_{i,j=1}^\infty |a_{ij}|^2$. Let $k_N(x,y) = \sum_{i,j=1}^N a_{ij} e_i(x) e_j(y)$. Define $K_N \in \mathcal{B}(X,Y)$ by

$$\begin{aligned}K_N h(x) &= \int_0^1 h(y) k_N(x,y)\,dy = \sum_{i,j=1}^N a_{ij} \int_0^1 h(y) e_i(x) e_j(y)\,dy \\ &= \sum_{i,j=1}^N a_{ij} \langle h, e_j \rangle e_i(x).\end{aligned}$$

This operator is finite rank, with $\operatorname{Ran} K_N \subset \operatorname{span}\{e_i : 1 \le i \le N\}$. By the first paragraph, $\|K - K_N\| = \|k - k_N\|_2 = \sum_{i>N \text{ or } j>N} |a_{ij}|^2 \to 0$ as $N \to \infty$. Therefore K is a norm limit of finite rank operators, and thus is compact.

5.1.5. EXAMPLE. Volterra Operator. Define an operator $V \in \mathcal{B}(L^2(0,1))$ by

$$Vh(x) = \int_0^x h(y)\,dy.$$

This is a Hilbert-Schmidt integral operator with kernel $k(x,y) = \begin{cases} 1 & \text{if } y \le x \\ 0 & \text{if } y > x \end{cases}$. Therefore $\|V\| \le \|k\|_2 = \dfrac{1}{\sqrt{2}}$.

First we show that V has no eigenvalues. Suppose that

$$\lambda h(x) = Vh(x) = \int_0^x h(y)\,dy.$$

If $\lambda = 0$, then $h = 0$ a.e. If $\lambda \ne 0$, then $h(x) = \lambda^{-1} \displaystyle\int_0^x h(y)\,dy$. The RHS is a continuous function, and hence $h \in C[0,1]$ with $h(0) = 0$. Then by the Fundamental Theorem of Calculus, the RHS is differentiable. Therefore $h'(x) = \lambda^{-1} h(x)$. This implies that $h(x) = ce^{x/\lambda}$. Since $h(0) = 0 = c$, we have $h = 0$. Therefore $\sigma_p(V) = \varnothing$.

Now we compute the powers of V.

$$V^2 h(x) = \int_0^x \int_0^y h(t)\,dt\,dy = \int_0^x \int_t^x dy\,h(t)\,dt = \int_0^x (x-t)h(t)\,dt.$$

$$V^3 h(x) = \int_0^x \int_0^y (y-t)h(t)\,dt\,dy = \int_0^x \int_t^x (y-t)\,dy\,h(t)\,dt = \int_0^x \frac{(x-t)^2}{2}\,h(t)\,dt.$$

We will show by induction that $V^n h(x) = \displaystyle\int_0^x \frac{(x-t)^{n-1}}{(n-1)!}\,h(t)\,dt$. This holds for $n = 1, 2, 3$ and assuming the formula for V^n,

$$\begin{aligned} V^{n+1} h(x) &= \int_0^x \int_0^y \frac{(y-t)^{n-1}}{(n-1)!}\,h(t)\,dt\,dy \\ &= \int_0^x \int_t^x \frac{(y-t)^{n-1}}{(n-1)!}\,dy\,h(t)\,dt = \int_0^x \frac{(y-t)^n}{n!}\,h(t)\,dt. \end{aligned}$$

Now from Example 5.1.4, we know that

$$\|V^n\| \le \Big\| \frac{(x-t)^{n-1}}{(n-1)!} \chi_{t \le x} \Big\|_2 \le \frac{1}{(n-1)!}.$$

Therefore $\operatorname{spr}(V) \le \lim_{n\to\infty} \big(\frac{1}{(n-1)!}\big)^{1/n} = 0$. Hence $\sigma(V) = \{0\}$. An operator with spectrum $\{0\}$ is called *quasinilpotent*.

5.1.6. PROPOSITION. *Let H be a Hilbert space. Then every $K \in \mathcal{K}(H)$ is a limit of finite rank operators.*

PROOF. Let $\varepsilon = \frac{1}{n}$. Choose an ε-net for $Kb_1(H)$, say $y_1, \dots, y_m$. Let P_n be the orthogonal projection onto span$\{y_1, \dots, y_m\}$. Then P_n is finite rank, and thus P_nK is also finite rank. Compute:

$$\|K - P_nK\| = \sup_{\|x\| \le 1} \|P_n^\perp Kx\| \le \sup_{\|x\| \le 1} \min_{i \le m} \|Kx - y_i\| \le \frac{1}{n}.$$

Therefore P_nK converge to K in norm. ∎

5.1.7. REMARK. This same argument will work in any Banach space X in which there is a sequence (or net) of finite rank projections (P_n) which are uniformly bounded and converge strongly to the identity, i.e., $P_nx \to x$ for every $x \in X$. This includes l_p for $1 \le p < \infty$. Here the projection P_n onto the span of the first n standard basis vectors does the job.

The existence of Banach spaces in which there are compact operators that are not limits of finite rank operators is a difficult result.

5.1.8. THEOREM (Schauder). *Let X, Y be Banach spaces. Then $T \in \mathcal{B}(X, Y)$ is compact if and only if $T^* \in \mathcal{B}(Y^*, X^*)$ is compact.*

PROOF. Assume that T is compact, and let $K = \overline{Tb_1(X)}$. This is a compact metric space. Define $R : \overline{b_1(Y^*)} \to C(K)$ be the restriction map $\psi \to \psi|_K$. If $\psi \in \overline{b_1(Y^*)}$, $|\psi(y_1) - \psi(y_2)| \le \|y_1 - y_2\|$. So the functions $R\psi$ are bounded (by $\|T\|$) and equicontinuous (all have Lipschitz constant 1). By the Arzela-Ascoli Theorem, $\overline{Rb_1(Y^*)}$ is compact in $C(K)$.

Therefore if $\{\psi_n : n \ge 1\}$ is any sequence in $\overline{b_1(Y^*)}$, there is a subsequence ψ_{n_i} such that $R\psi_{n_i}$ converge uniformly. That is, $\psi_{n_i}|_K$ converge uniformly to a function $\Psi \in C(K)$. This implies $\psi_{n_i}(Tx) = T^*\psi_{n_i}(x)$ converges uniformly to $\Psi(Tx) = \Psi \circ T(x)$ for $x \in b_1(X)$. Hence $\varphi = \Psi \circ T \in X^*$ and $T^*\psi_{n_i} \to \varphi$ uniformly over the unit ball; i.e. converges in norm. This shows that $T^*b_1(Y^*)$ is precompact; whence T^* is a compact operator.

The converse is now easy. If T^* is compact, then T^{**} is compact. Then considering X as a subspace of X^{**} and Y as a subspace of Y^{**}, we see that $\overline{Tb_1(X)} \subset \overline{T^{**}b_1(X^{**})}$. This is a closed subset of a compact set, and hence is compact. Therefore T is a compact operator. ∎

5.1.9. REMARK. For the Hilbert space adjoint, there is an easier proof based on Proposition 5.1.6. If T is compact in $\mathcal{B}(H)$, then $T = \lim_{n\to\infty} F_n$ where F_n are finite rank. Then $T^* = \lim_{n\to\infty} F_n^*$ is also a limit of finite rank operators, and thus is compact.

5.2. Structure of Compact Operators

5.2.1. LEMMA. *Let X be a Banach space, and let V be a closed subspace such that either V or* $\dim X/V$ *is finite dimensional. Then V is complemented.*

PROOF. If $\dim V = n < \infty$, let $v_1, \dots, v_n$ be a basis for V. There is a dual basis $\varphi_1, \dots, \varphi_n$ for V^* given by $\varphi_i(v_j) = \delta_{ij}$. Using the Hahn-Banach Theorem, we can extend each φ_j to some $\tilde{\varphi}_j \in X^*$. Then $P = \sum_{i=1}^n v_i \varphi_i \in \mathcal{B}(X)$. Observe that $Px = \sum_{i=1}^n v_i \varphi_i(x) \in V$ and $Pv_j = v_j$. Therefore P is a continuous idempotent with range V. Hence $X = V \dotplus \ker P$.

If $\dim X/V = n < \infty$, let $\dot{x}_1, \dots, \dot{x}_n$ be a basis for X/V. Choose $x_i \in X$ so that $\dot{x}_i = x_i + V$. Let $W = \operatorname{span}\{x_1, \dots, x_n\}$. Then $V + W = X$ because if $x \in X$, then $\dot{x} := x + V = \sum_{i=1}^n a_i \dot{x}_i$. Therefore $x - \sum_{i=1}^n a_i x_i \in V$. Also if $x \in V \cap W$, write $x = \sum_{i=1}^n a_i x_i$ then $\dot{x} = \sum_{i=1}^n a_i \dot{x}_i = 0$; so that all $a_i = 0$ and so $x = 0$. Therefore $X = V \dotplus W$. ■

5.2.2. KEY LEMMA. *Let $K \in \mathcal{K}(X)$. Suppose that there are closed subspaces $V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \dots$ and scalars $\alpha_i \in \mathbb{C}$ so that $(\alpha_i I - K)V_i \subset V_{i-1}$ for $i \geq 1$. Then* $\lim_{i\to\infty} \alpha_i = 0$.

PROOF. Choose unit vectors $x_i \in V_i$ with $\operatorname{dist}(x_i, V_{i-1}) \geq \frac{1}{2}$ for $i \geq 1$. Then $Kx_i = \alpha_i x_i + y_i$ where $y_i \in V_{i-1}$. Suppose that there is a subsequence with $|\alpha_{i_n}| \geq \delta > 0$ for $i_1 < i_2 < \dots$. Then if $m < n$, since $y_{i_n} - Kx_{i_m} \in V_{i_n - 1}$,

$$\|Kx_{i_n} - Kx_{i_m}\| = \|\alpha_{i_n} x_{i_n} + (y_{i_n} - Kx_{i_m})\| \geq \frac{1}{2}|\alpha_{i_n}| \geq \frac{\delta}{2}.$$

This shows that $\overline{Kb_1(X)}$ is not compact, a contradiction. ■

Write $\operatorname{null}(T) = \dim \ker T$ for the *nullity* of $T \in \mathcal{B}(X)$.

5.2.3. THEOREM. *Let $K \in \mathcal{K}(X)$. Then for $\lambda \neq 0$,* $\ker(\lambda I - K)$ *is finite dimensional and* $\operatorname{Ran}(\lambda I - K)$ *is closed and finite codimension.*

PROOF. By replacing $\lambda I - K$ by $I - \lambda^{-1}K$ and K by $\lambda^{-1}K$, we may suppose that $\lambda = 1$. Let $B = \overline{b_1(X)} \cap \ker(I - K)$. Then the unit ball of $N = \ker(I - K)$ is $B = KB \subset \overline{Kb_1(X)}$ which is compact. Therefore $\operatorname{null}(I - K) < \infty$ by Proposition 2.2.15.

Choose a complement for N so that $X = N \dotplus V$. Then $I - K|_V$ is injective and $\operatorname{Ran}(I - K) = (I - K)V$. Claim: $(I - K)|_V$ is bounded below. If not, there are unit vectors $v_n \in V$ so that $\|(I - K)v_n\| \to 0$. By compactness, there is a subsequence v_{n_i} so that $Kv_{n_i} \to y$. Since $(I - K)v_{n_i} \to 0$, we get that $v_{n_i} \to y$. Therefore $\|y\| = \lim_{i\to\infty} \|v_{n_i}\| = 1$ and $(I - K)y = \lim_{i\to\infty}(I - K)v_{n_i} = 0$. So

$y \in V \cap N$, which is impossible. Hence $I - K$ is bounded below on V, and thus $R = (I-K)V = \operatorname{Ran}(I-K)$ is closed.

If X/R is infinite dimensional, choose vectors $x_i \in X$ so that $\dot{x}_i = x_i + R$ are linearly independent for $i \geq 1$. Define $R_n = \operatorname{span}\{R, x_1, \dots, x_n\}$. Then R_n/R is finite dimensional, and thus closed in X/R; so that $R_n = Q^{-1}(R_n/R)$ is closed because it is the preimage of a closed subspace by the quotient map. Clearly

$$R \subsetneq R_1 \subsetneq R_n \subsetneq R_{n+1} \subsetneq \dots$$

However $(I-K)R_n \subset \operatorname{Ran}(I-K) = R$. By the Key Lemma, we have that $\lim_{n\to\infty} 1 = 0$ which is absurd. Therefore R has finite codimension. ∎

5.2.4. DEFINITION. If $\mathcal{S} \subset \mathcal{B}(X)$, the *commutant* of $\mathcal{S}$ is

$$\mathcal{S}' = \{T \in \mathcal{B}(X) : TS = ST \text{ for all } S \in \mathcal{S}\}.$$

We write $\mathcal{S}''$ for $(\mathcal{S}')'$.

Note that $\mathcal{S}'$ is an algebra (a vector space closed under multiplication) containing the identity. Moreover it is closed in the weak operator topology.

5.2.5. THEOREM. *Let $K \in \mathcal{K}(X)$. Then for $\lambda \neq 0$, there is an integer n_0 so that*

$$N(\lambda) := \ker(\lambda I - K)^{n_0} = \ker(\lambda I - K)^{n_0+k}$$

and

$$R(\lambda) := \operatorname{Ran}(\lambda I - K)^{n_0} = \operatorname{Ran}(\lambda I - K)^{n_0+k}$$

for $k \geq 1$. Moreover $X = N(\lambda) \dotplus R(\lambda)$. The subspaces $N_n = \ker(\lambda I - K)^n$ and $R_n = \operatorname{Ran}(\lambda I - K)^n$ are invariant for $\{K\}'$ and

$$\operatorname{null}(\lambda I - K)^n = \dim X/\operatorname{Ran}(\lambda I - K)^n = \operatorname{null}(\lambda I - K^*)^n \quad \textit{for} \quad n \geq 1.$$

PROOF. Let $N_n = \ker(\lambda I - K)^n$ for $n \geq 0$. Since $(\lambda I - K)^n = \lambda^n I - K_n$ where K_n is compact, these are finite dimensional subspaces. Note that $N_{n-1} \subset N_n$ and $(\lambda I - K)N_n \subset N_{n-1}$ for $n \geq 1$. If $N_n \supsetneq N_{n-1}$ for all $n \geq 1$, the Key Lemma would imply that $\lim_{n\to\infty} \lambda = 0$, which is absurd. Therefore for some smallest $n_0 \geq 0$, $N_{n_0} = N_{n_0+1}$. We show by induction that $N_{n_0} = N_{n_0+k}$ for $k \geq 1$. It is valid for $k = 1$. Suppose that it holds for k. If $x \in N_{n_0+k+1}$, then $(\lambda I - K)x \in N_{n_0+k} = N_{n_0}$. Therefore $x \in N_{n_0+1} = N_{n_0}$. So $N_{n_0+k+1} = N_{n_0}$.

Now let $R_n = \operatorname{Ran}((\lambda I - K)^n)$ for $n \geq 0$. Since $(\lambda I - K)^n = \lambda^n I - K_n$ where K_n is compact, these are closed subspaces of finite codimension. Note that $R_{n-1} \supset R_n$ and $(\lambda I - K)R_{n-1} = R_n$ for $n \geq 1$. Therefore by Lemma 4.1.7, $N_n^* := R_n^\perp = \ker(\lambda I - K^*)^n$ are finite dimensional subspaces of X^* such that $N_{n-1}^* \subset N_n^*$ and $(\lambda I - K^*)N_n^* \subset N_{n-1}^*$. By Schauder's Theorem, K^* is compact. So arguing as in the previous paragraph, there is a least integer m_0 so that $N_{m_0+k}^* = N_{m_0}^*$ and thus $R_{m_0+k} = R_{m_0}$ for all $k \geq 1$.

Let $p = \max\{m_0, n_0\}$ (they turn out to be equal). Note that if $T \in \{K\}'$ and $x \in N_k$, then $0 = T(\lambda I - K)^k x = (\lambda I - K)^k Tx$; whence $TN_k \subset N_k$. Also if $y = (\lambda I - K)^k x \in R_k$, then $Ty = T(\lambda I - K)^k x = (\lambda I - K)^k Tx$; whence $TR_k \subset R_k$. So N_k and R_k are $\{K\}'$ invariant subspaces for each $k \geq 1$.

Let $x \in X$. Since $(\lambda I - K)^p x \in R_p = R_{2p}$, there is a vector $y \in X$ so that $(\lambda I - K)^p x = (\lambda I - K)^{2p} y$. Let $z = x - (\lambda I - K)^p y$. Observe that $(\lambda I - K)^p z = (\lambda I - K)^p x - (\lambda I - K)^{2p} y = 0$; whence $z \in N_p$. Therefore $x = z + (\lambda I - K)^p y \in N_p + R_p$. On the other hand, if $x \in N_p \cap R_p$, there is $y \in X$ so that $(\lambda I - K)^p y = x$. Therefore $(\lambda I - K)^{2p} y = (\lambda I - K)^p x = 0$, so that $y \in N_{2p} = N_p$. Hence $x = 0$. So $N_p \cap R_p = \{0\}$. Therefore $X = N_p \dotplus R_p$.

Next observe that $R_p = (\lambda I - K)^p N_p + (\lambda I - K)^p R_p = (\lambda I - K)^p R_p$. Therefore $(\lambda I - K)^p$ maps R_p one to one and onto itself. That is, $(\lambda I - K)^p|_{R_p}$ is invertible. Since $\lambda I - K$ maps R_p into itself, it is also bijective on R_p. Therefore $\lambda I - K = (\lambda I - K)|_{N_p} \dotplus (\lambda I - K)|_{R_p}$, and $F_\lambda = (\lambda I - K)|_{N_p}$ acts on a finite dimensional space, and is nilpotent of order at most p. It follows that $\sigma(F_\lambda) = \{0\}$ and so $\sigma(K|_{N_p}) = \{\lambda\}$.

Since $(\lambda I - K)|_{R_p}$ is invertible, we have that

$$\ker(\lambda I - K)^n = \ker(\lambda I - K)^n|_{N_p}$$

and

$$X/\operatorname{Ran}(\lambda I - K)^n = N_p/\operatorname{Ran}(\lambda I - K)^n|_{N_p}.$$

For a linear map A on a finite dimensional space N, we have $\operatorname{null}(A) + \dim \operatorname{Ran} A = \dim N$. Therefore $\operatorname{null}(A) = \dim N/\operatorname{Ran} A = \operatorname{null}(A^*)$. In particular

$$\operatorname{null}(\lambda I - K)^n = \dim X/\operatorname{Ran}(\lambda I - K)^n = \operatorname{null}(\lambda I - K^*)^n \quad \text{for} \quad n \geq 1.$$

In particular, we have $m_0 = n_0$. ■

5.2.6. REMARK. Since $(\lambda I - K)|_N(\lambda)$ is nilpotent on a finite dimensional space, we have $\sigma(K|_{N(\lambda)}) = \{\lambda\}$. By a familiar result in linear algebra, this restriction is similar to a sum of Jordan blocks for the eigenvalue λ.

This theorem was the main technical result. We put everything together in the following.

5.2.7. STRUCTURE OF COMPACT OPERATORS. *Let X be an infinite dimensional Banach space, and let $K \in \mathcal{K}(X)$. Then*

(1) $0 \in \sigma(K)$. *If* $0 \neq \lambda \in \sigma(K)$, *then* $\lambda \in \sigma_p(K)$.

(2) *The spectrum is finite or is a countable set* $\{0, \lambda_n : n \geq 1\}$ *such that* $\lim_{n\to\infty} \lambda_n = 0$.

(3) *For each* $\lambda \in \sigma(K) \setminus \{0\}$, *there is an integer* n_λ *so that*

$$N(\lambda) = \ker(\lambda I - K)^{n_\lambda} = \ker(\lambda I - K)^{n_\lambda + k} \quad \textit{for} \quad k \geq 1,$$

$$R(\lambda) = \mathrm{Ran}(\lambda I - K)^{n_\lambda} = \mathrm{Ran}(\lambda I - K)^{n_\lambda + k} \quad \textit{for} \quad k \ge 1,$$

$N(\lambda)$ is finite dimensional, $R(\lambda)$ is closed and $X = N(\lambda) \dotplus R(\lambda)$.

(4) *$\sigma(K|_{N(\lambda)}) = \{\lambda\}$ and $\sigma(K|_{R(\lambda)}) = \sigma(K) \setminus \{\lambda\}$.*

(5) *There is a unique finite rank idempotent E_λ in $\{K\}''$ with range $N(\lambda)$ and kernel $R(\lambda) = \mathrm{Ran}(\lambda I - K)^{n_\lambda}$, both invariant for $\{K\}'$.*

(6) *If $\lambda, \mu \in \sigma(K) \setminus \{0\}$ are distinct, then $E_\lambda E_\mu = 0$.*

PROOF. Compact operators can never be surjective because the Open Mapping Theorem would then show that $Kb_1(X)$ contains a ball $b_r(X)$. So $0 \in \sigma(K)$. Theorem 5.2.5 shows that if $\lambda \neq 0$, then either $\ker(\lambda I - K) \neq \{0\}$ or $\lambda I - K$ is invertible. Thus $\sigma(K) = \sigma_p(K) \cup \{0\}$. If $\lambda \in \sigma(K) \setminus \{0\}$, then $X = N(\lambda) \dotplus R(\lambda)$ where both $N(\lambda)$ and $R(\lambda)$ are $\{K\}'$-invariant. By Lemma 4.4.8,

$$\sigma(K) = \sigma(K|_{N(\lambda)}) \cup \sigma(K|_{R(\lambda)}) = \{\lambda\} \cup \sigma(K|_{R(\lambda)}).$$

Since $(\lambda I - K)|_{R(\lambda)}$ is invertible, $\sigma(K|_{R(\lambda)}) = \sigma(K) \setminus \{\lambda\}$. Therefore λ is an isolated point of the spectrum, and hence the spectrum is at most countable. If $\sigma(K) \setminus \{0\} = \{\lambda_n : n \ge 1\}$ is countable, pick a unit vector $x_n \in \ker(\lambda_n I - K)$. Define $V_n = \mathrm{span}\{x_1, \dots, x_n\}$. Observe that $V_{n-1} \subsetneq V_n$ and $(\lambda_n - K)V_n \subset V_{n-1}$ for $n \ge 1$. By the Key Lemma, $\lim_{n\to\infty} \lambda_n = 0$.

For (5), there is a unique idempotent with range $N(\lambda)$ and kernel $R(\lambda)$ by Section 2.7. If $T \in \{K\}'$, then T leaves the range and kernel of E_λ invariant. That means that

$$TE_\lambda = E_\lambda T E_\lambda \quad \text{and} \quad T(I - E_\lambda) = (I - E_\lambda)T(I - E_\lambda).$$

This implies that $TE_\lambda = E_\lambda T E_\lambda = E_\lambda T$. Thus $E_\lambda \in \{K\}''$.

Suppose that μ is another non-zero point in $\sigma(K)$. Then $E_\mu \in \{K\}'' \subset \{K\}'$, and hence $E_\mu E_\lambda = E_\lambda E_\mu$. Therefore $\mathrm{Ran}(E_\mu E_\lambda) \subset N(\lambda) \cap N(\mu) = \{0\}$ because $(\lambda I - K)^{n(\lambda)}$ annihilates $N(\lambda)$ while $(\mu I - K)^{n(\mu)}$ is invertible on $N(\lambda)$. Therefore $E_\lambda E_\mu = 0$. ■

5.3. Fredholm Operators

5.3.1. DEFINITION. An operator $T \in \mathcal{B}(X, Y)$ is *Fredholm* if $\mathrm{null}(T) < \infty$, $\mathrm{Ran}\, T$ is closed and has finite codimension. The *Fredholm index* of T is

$$\mathrm{ind}\, T = \mathrm{null}(T) - \dim(Y/TX) = \mathrm{null}(T) - \mathrm{null}(T^*).$$

5.3.2. Remarks. The fact that $\operatorname{Ran} T$ is closed is automatic. Suppose that $Y = \operatorname{span}\{TX, y_1, \dots, y_m\}$ where $\dim(Y/TX) = m$. Define

$$S : X/\ker T \oplus \mathbb{C}^m \to Y \quad \text{by} \quad S(\dot{x}, a) = Tx + \sum_{i=1}^{m} a_i y_i.$$

This is a continuous bijection, and hence an isomorphism. Therefore $TX = S(X/\ker T)$ is closed.

5.3.3. Examples.
(1) Every invertible operator is Fredholm of index 0.

(2) The unilateral shift $S \in \mathcal{B}(l_p)$ for $1 \le p < \infty$ is an isometry, and thus has no kernel and has closed range. Since $l_p/Sl_p \simeq \mathbb{C}$, we see that S is Fredholm and $\operatorname{ind} S = -1$. Similarly if T is the backward shift, then $Tl_p = l_p$ and $\ker T = \mathbb{C}e_1$. So T is Fredholm and $\operatorname{ind} T = 1$. Recall that $TS = I$ and $ST = I - e_1\delta_1$ both lie in $I + \mathcal{K}(l_p)$. (Here $\delta_i(a_n) = a_i$ and $(e_1\delta_1)(x) = \delta_1(x)e_1$.)

5.3.4. Proposition. *If $\lambda \ne 0$ and $K \in \mathcal{K}(X)$, then $\lambda I - K$ is Fredholm and* $\operatorname{ind}(\lambda I - K) = 0$.

Proof. By Theorem 5.2.5, $\lambda I - K$ has closed range of finite codimension, and $\operatorname{null}(\lambda I - K) = \operatorname{null}(\lambda I - K^*)$. Hence

$$\operatorname{ind}(\lambda I - K) = \operatorname{null}(\lambda I - K) - \operatorname{null}(\lambda I - K^*) = 0. \qquad \blacksquare$$

5.3.5. Theorem. *The set $\mathcal{F}(X, Y)$ of Fredholm operators in $\mathcal{B}(X, Y)$ is open. Index is a continuous integer valued function, and hence is locally constant.*

Proof. Let T be Fredholm. Set $N = \ker T$ and choose a closed complement V; so that $X = N \dotplus V$. Then $TV = TX$. Since TX is closed and has finite codimension, choose a finite dimensional subspace W so that $Y = TX \dotplus W$. Define $\tilde{T} : V \oplus_1 W \to Y$ by $\tilde{T}(v, w) = Tv + w$. Note that

$$\|\tilde{T}(v, w)\| \le \|Tv\| + \|w\| \le \max\{\|T\|, 1\}(\|v\| + \|w\|).$$

So $\tilde{T}$ is continuous. It is easy to check that $\tilde{T}$ is a bijection, and hence it is invertible by the Banach Isomorphism Theorem.

Now suppose that $S \in \mathcal{B}(X, Y)$ such that $\|S - T\| < 1/\|\tilde{T}^{-1}\|$. Define $\tilde{S} : V \oplus W \to Y$ by $\tilde{S}(v, w) = Sv + w$. Then $\tilde{S}$ is continuous and

$$\|\tilde{S} - \tilde{T}\| = \|(S - T)|_V\| < 1/\|\tilde{T}^{-1}\|.$$

Therefore $\tilde{S}$ is invertible. It follows that $SV = \tilde{S}V$ is closed and

$$Y = \tilde{S}(V \oplus W) = SV \dotplus W.$$

Therefore $\dim Y/SX \le \dim Y/SV = \dim W < \infty$. Also $\ker S \cap V = \{0\}$, and so the quotient $Q : X \to X/V$ is injective on $\ker S$; so that

$$\operatorname{null} S \le \dim X/V = \dim N < \infty.$$

Therefore S is Fredholm.

The subspace $V \dotplus \ker S = Q^{-1}((\ker S + V)/V)$ is closed and finite codimension. Choose a finite dimensional complement Z so that $X = V \dotplus \ker S \dotplus Z = \ker S \dotplus (V \dotplus Z)$. Then $SX = SV \dotplus SZ$ because S is injective on $V \dotplus Z$ and Z is finite dimensional. Let $P : Y = SV \dotplus W \to W$ be the projection onto W with kernel SV. Then $PSX = PSZ \simeq Z$ because S is injective on Z and P is injective on SZ. Therefore

$$Y/SX = (Y/SV)/(SX/SV) \simeq W/PSZ.$$

We compute

$$\begin{aligned} \operatorname{ind} S &= \operatorname{null} S - \dim Y/SX \\ &= \operatorname{null} S - (\dim W - \dim Z) \\ &= \dim(\ker S \dotplus Z) - \dim W \\ &= \operatorname{null} T - \dim Y/TX = \operatorname{ind} T. \end{aligned}$$

Thus index is constant on an open ball around T, so that it is locally constant, whence continuous. ∎

5.3.6. COROLLARY. *Fredholm index is constant on connected components of* $\mathcal{F}(X, Y)$.

We also get the following result which is a consequence of our proof. We showed that if $\|S - T\| < 1/\|\tilde{T}^{-1}\|$, then $\operatorname{null} S \le \operatorname{null} T$.

5.3.7. COROLLARY. *If* $T \in \mathcal{B}(X, Y)$ *is Fredholm, then* $\limsup_{S \to T} \operatorname{null} S \le \operatorname{null} T$.

Let $\pi : \mathcal{B}(X) \to \mathcal{B}(X)/\mathcal{K}(X)$ be the quotient map. This is a Banach space, and since $\mathcal{K}(X)$ is an ideal, it is also a ring. Multiplication is continuous: if $a, b \in \mathcal{B}(X)/\mathcal{K}(X)$ then we can choose $A, B \in \mathcal{B}(X)$ so that $\pi(A) = a$ and $\|A\| < (1 + \varepsilon)\|a\|$ and $\pi(B) = b$ and $\|B\| < (1 + \varepsilon)\|b\|$. Then

$$\|ab\| = \|\pi(AB)\| \le \|A\|\,\|B\| < (1 + \varepsilon)^2 \|a\|\,\|b\|.$$

This is called a *Banach algebra* when it is a Banach space and a ring in which multiplication is continuous. In particular, we can talk about invertible elements in $\mathcal{B}(X)/\mathcal{K}(X)$.

5.3.8. ATKINSON'S THEOREM. $T \in \mathcal{B}(X)$ *is Fredholm if and only if* $\pi(T)$ *is invertible in* $\mathcal{B}(X)/\mathcal{K}(X)$.

PROOF. Suppose that T is Fredholm. Let $N = \ker T$ and let V be a complement for N, so that $X = N \dot{+} V$. Also $TX = TV$ is closed with finite codimension, and we pick a (finite dimensional) complement W so that $X = TV \dot{+} W$. If we consider $T|_V$ as an element of $\mathcal{B}(V, TV)$, then this is a continuous bijection. Hence by Banach's Isomorphism Theorem, it has a continuous inverse $S \in \mathcal{B}(TV, V)$. Define $\tilde{S} \in \mathcal{B}(X)$ by $\tilde{S}(Tv + w) = v$ for $v \in V$ and $w \in W$. Note that $\tilde{S}T(u + v) = \tilde{S}Tv = v$ for $u \in N$ and $v \in V$. This is $\tilde{S}T = P_V$ is the projectiion of X onto V with kernel N, Now $I - P_V = P_N$ is finite rank. Therefore $\pi(\tilde{S})\pi(T) = \pi(I - P_N) = \pi(I)$. Similarly, $T\tilde{S}(Tv + w) = Tv = P_{TV}(Tv + w)$ is the projection onto TV with kernel W. Again $I - P_{TV} = P_W$ is finite rank, so that $\pi(T)\pi(\tilde{S}) = \pi(I)$. Therefore $\pi(T)$ is invertible.

Conversely, suppose that $\pi(T)$ has an inverse $\pi(S)$ in $\mathcal{B}(X)/\mathcal{K}(X)$. Then there are compact operators $K, L \in \mathcal{K}(X)$ so that $ST = I + K$ and $TS = I + L$. Then $\ker T \subseteq \ker I + K$ is finite dimensional and $\operatorname{Ran} T \supseteq \operatorname{Ran}(I + L)$, so it has finite codimension. Therefore T is Fredholm. ■

5.3.9. COROLLARY. *If $T \in \mathcal{B}(X)$ is Fredholm and $K \in \mathcal{K}(X)$, then $T + K$ is Fredholm and* $\operatorname{ind}(T + K) = \operatorname{ind}(T)$.

PROOF. By Atkinson's Theorem, $\pi(T + K) = \pi(T)$ is invertible and hence $T + K$ is Fredholm. Thus $T + tK$ for $0 \leq t \leq 1$ is a continuous path of Fredholm operators. Since index is locally constant, we have $\operatorname{ind}(T + K) = \operatorname{ind}(T)$. ■

It follows that the following definition is well defined.

5.3.10. DEFINITION. If $a \in \mathcal{B}(X)/\mathcal{K}(X)$ is invertible, define $\operatorname{ind}(a) = \operatorname{ind}(T)$ for any T with $\pi(T) = a$.

5.3.11. THEOREM. *If $T \in \mathcal{B}(X, Y)$ and $S \in \mathcal{B}(Y, Z)$ are both Fredholm, then ST is Fredholm and* $\operatorname{ind}(ST) = \operatorname{ind}(S) + \operatorname{ind}(T)$.

PROOF. As before, let $N_S = \ker S$ and choose a complement $Y_0 \subset Y$ so that $Y = N_S \dot{+} Y_0$. Also choose a finite dimensional complement Z_0 for SY so that $Z = SY \dot{+} Z_0$. Choose a complement X_0 for $N_T = \ker T$ so that $X = N_T \dot{+} X_0$. Finally we reqiure a complement W for TX in Y, but we do this carefully in two steps. First choose a complement $W_0 \subset N_S$ for $N_S \cap TX$ in N_S so that $N_S = (N_S \cap TX) \dot{+} W_0$. Note that $W_0 \cap TX \subset W_0 \cap (N_S \cap TX) = \{0\}$, so that $TX + W_0 = TX \dot{+} W_0 = TX + N_S$. Now choose a complement W_1 for $TX + N_S$, so that $Y = TX \dot{+} W_0 \dot{+} W_1$.

Recall that $TX = TV$ and $T|_V$ maps V bijectively onto TX. We compute

$$\ker ST = \ker T + \{v \in V : Tv \in N_S\} = N_T \dot{+} (T|_V)^{-1}(N_S \cap TX).$$

Therefore null$(ST) =$ null$(T) +$ dim$(N_S \cap TX)$. Next,

$$SY = S(TX \dotplus W_0 \dotplus W_1) = STX \dotplus SW_1.$$

The latter sum is a direct sum because if $w_1 \in W_1$ and $Sw_1 = STx$, then $w_1 - Tx$ lies in N_S, and hence $w_1 \in W_1 \cap (N_S + TX) = \{0\}$. Moreover, W_1 does not intersect N_S, so that S is injective on W_1. Therefore

$$\dim Z/SY = \dim Z/(STX \dotplus SW_1) = \dim Z/STX - \dim W_1.$$

Finally we can calculate

$$\begin{aligned}
\operatorname{ind} ST &= \operatorname{null} ST - \dim Z/STX \\
&= \operatorname{null} T + \dim(N_S \cap TX) - \big(\dim Z/SY + \dim W_1\big) \\
&= \operatorname{null} T + \big(\dim(N_S \cap TX) + \dim W_0\big) - \dim(W_0 \dotplus W_1) - \dim Z/SY \\
&= \operatorname{null} T + \operatorname{null} S - \dim Y/TX - \dim Z/SY = \operatorname{ind} S + \operatorname{ind} T. \qquad \blacksquare
\end{aligned}$$

The invertible elements of any ring form a group under multiplication. When we restate Theorem 5.3.11 as a result about $\mathcal{B}(X)/\mathcal{K}(X)$, we obtain:

5.3.12. COROLLARY. *The Fredholm index is a homomorphism from the group* $(\mathcal{B}(X)/\mathcal{K}(X))^{-1}$ *into* $(\mathbb{Z}, +)$.

5.4. Normal Operators

5.4.1. DEFINITION. Let H be a Hilbert space. An operator $T \in \mathcal{B}(H)$ is *self-adjoint* if $T^* = T$. It is *positive* if $T^* = T$ and $\langle Tx, x\rangle \geq 0$ for $x \in H$. It is *unitary* if $T^* = T^{-1}$. And it is *normal* if $T^*T = TT^*$.

5.4.2. REMARKS.
(1) If $T : H \to H$ is everywhere defined and formally self-adjoint, meaning that $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$, then T is automatically bounded. See Example 2.4.13. This is known as the Hellinger-Toeplitz Theorem.

(2) If H is a complex Hilbert space and $T \in \mathcal{B}(H)$ satisfies $\langle Tx, x\rangle \geq 0$ for $x \in H$, then the polarization identity shows that

$$\langle Tx, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \langle T(x + i^k y), x + i^k y\rangle$$

and

$$\langle T^*x, y\rangle = \langle x, Ty\rangle = \overline{\langle Ty, x\rangle} = \overline{\frac{1}{4}\sum_{k=0}^{3} i^k \langle T(y + i^k x), y + i^k x\rangle}$$

$$= \frac{1}{4}\sum_{k=0}^{3}(-i)^k\langle T(y+i^k x), y+i^k x\rangle$$

$$= \frac{1}{4}\sum_{k=0}^{3}(-i)^k\langle T(x+(-i)^k y), x+(-i)^k y\rangle = \langle Tx, y\rangle$$

Hence $T^* = T$ is automatically self-adjoint.

If H is a real Hilbert space, then $T = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] \in \mathcal{M}_2(\mathbb{R})$ satisfies

$$\left\langle \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = -yx + xy = 0$$

for all $(x, y) \in \mathbb{R}^2$. However this matrix represents a 90 degree rotation, and is clearly not self-adjoint because $T^* = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] = -T$ even though $\langle Tx, x\rangle \geq 0$ for all $x \in \mathbb{R}^2$.

(3) If U is unitary, then $\langle Ux, Uy\rangle = \langle U^*Ux, y\rangle = \langle x, y\rangle$. Thus unitary maps are invertible isometries and they preserve the inner product.

5.4.3. Examples.

(1) In Example 4.3.8(1), we showed that the multiplication operator M_f on $L^2(0,1)$ is bounded when $f \in L^\infty(0,1)$ and that $M_f^* = M_{\bar{f}}$. It is clear that this is a normal operator. It is self-adjoint if and only if $f = \bar{f}$ a.e., i.e., when f is real valued. It is positive if and only if $f \geq 0$ a.e.. And it is unitary if and only if $|f|^2 = 1$ a.e.

A similar analysis applies to $D = \operatorname{diag}(d_1, d_2, d_3,]\dots) \in \mathcal{B}(l_2)$.

(2) If $A \in \mathcal{B}(H)$, then A^*A is positive. Indeed, $(A^*A)^* = A^*A^{**} = A^*A$ and $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0$. Moreover $\|A^*A\| \leq \|A^*\|\,\|A\| = \|A\|^2$ and

$$\|A^*A\| = \sup_{\|x\|\leq 1}\|A^*Ax\| \geq \sup_{\|x\|\leq 1}\langle A^*Ax, x\rangle = \sup_{\|x\|\leq 1}\|Ax\|^2 = \|A\|^2.$$

Hence $\|A^*A\| = \|A\|^2$.

5.4.4. Proposition. *Let $N \in \mathcal{B}(H)$ be a normal operator. Then*

(1) $\|Nx\| = \|N^*x\|$ *for all* $x \in H$.

(2) $\|N\| = \operatorname{spr}(N)$.

(3) $\ker(\lambda I - N) = \ker(\lambda I - N)^n = \ker(\lambda I - N)^*$ *for all* $\lambda \in \mathbb{C}$ *and* $n \geq 1$.

(4) $\ker(\lambda I - N)^\perp = \overline{\operatorname{Ran}(\lambda I - N)}$ *for all* $\lambda \in \mathbb{C}$.

(5) $\ker(\lambda I - N) \perp \ker(\mu I - N)$ *if* $\lambda \neq \mu \in \mathbb{C}$.

(6) *If* $p \in \mathbb{C}[z]$ *is a polynomial, then* $\|p(N)\| = \sup\big\{|p(\lambda)| : \lambda \in \sigma(N)\big\}$.

PROOF. $$\|Nx\|^2 = \langle Nx, Nx\rangle = \langle N^*Nx, x\rangle = \langle NN^*x, x\rangle = \langle N^*x, N^*x\rangle = \|N^*x\|^2.$$

So (1) holds. For (2), $\|N^2x\| = \|N^*(Nx)\| \geq \langle N^*Nx, x\rangle = \|Nx\|^2$. Taking the supremum over $\|x\| \leq 1$ yields $\|N^2\| \geq \|N\|^2$, while $\|N^2\| \leq \|N\|^2$ is true for all operators. Thus $\|N^2\| = \|N\|^2$. Iterating this yields $\|N^{2^k}\| = \|N\|^{2^k}$ for $k \geq 1$. Therefore

$$\operatorname{spr}(N) = \lim_{k\to\infty} \|N^{2^k}\|^{1/2^k} = \|N\|.$$

Since $\lambda I - N$ is normal, we have by (1) that $\|(\lambda I - N)x\| = \|(\lambda I - N)^*x\|$. In particular, $x \in \ker(\lambda I - N)$ if and only if $x \in \ker(\lambda I - N)^*$. Also by the proof of (2), $\|(\lambda I - N)^{2^k}x\| \geq \|(\lambda I - N)x\|^{2^k}$. Hence $x \in \ker(\lambda I - N)^{2^k}$ implies that $x \in \ker(\lambda I - N)$. The other inclusion is trivial. So $\ker(\lambda I - N)^{2^k} = \ker(\lambda I - N)$. Hence $\ker(\lambda I - N)^n = \ker(\lambda I - N)$ for all $n \geq 1$. So (3) holds.

(4) $\overline{\operatorname{Ran}(\lambda I - N)} = \big(\ker(\lambda I - N)^*\big)^\perp = \big(\ker(\lambda I - N)\big)^\perp$.

(5) If $x \in \ker(\lambda I - N)$ and $y \in \ker(\mu I - N)$, then $Nx = \lambda x$ and $Ny = \mu y$, whence $N^*y = \overline{\mu}y$. Hence

$$\lambda\langle x, y\rangle = \langle Nx, y\rangle = \langle x, N^*y\rangle = \langle x, \overline{\mu}y\rangle = \mu\langle x, y\rangle.$$

Since $\lambda \neq \mu$, this forces $\langle x, y\rangle = 0$. Therefore $\ker(\lambda I - N) \perp \ker(\mu I - N)$.

(6) It is easy to check that $p(N)$ is normal. By the Spectral Mapping Theorem, $\sigma(p(N)) = p(\sigma(N))$. Therefore

$$\|p(N)\| = \operatorname{spr}(p(N)) = \sup\{|\mu| : \mu \in \sigma(p(N))\} = \sup\{|p(\lambda)| : \lambda \in \sigma(N)\}. \blacksquare$$

5.4.5. COROLLARY. *If a normal operator N is Fredholm, then* $\operatorname{ind} N = 0$.

PROOF. Compute: $\operatorname{ind} N = \operatorname{null} N - \operatorname{null} N^* = 0$. ■

5.4.6. SPECTRAL THEOREM FOR COMPACT NORMAL OPERATORS. *Suppose that $K \in \mathcal{B}(H)$ is compact and normal. Let $\sigma(K) = \{\lambda_n : n \geq 0\}$ where $\lambda_0 = 0 = \lim_{n\to\infty} \lambda_n$. Let $M_n = \ker(\lambda_n I - K)$ for $n \geq 0$. These are finite dimensional for $n \geq 1$ and $M_n \perp M_n$ for $m \neq n$. Moreover $H = \sum_{n\geq 0}^{\oplus} M_n$ and $K = \sum_{n\geq 0} \lambda_n P_{M_n}$ is diagonalizable by an orthonormal basis.*

PROOF. By the Structure Theorem for Compact Operators 5.2.7, $\sigma(K)$ contains 0, and is either finite or countable, $\{\lambda_n : n \geq 0\}$, where $\lambda_0 = 0 = \lim_{n\to\infty} \lambda_n$. For each $0 \neq \lambda_n \in \sigma(K) \setminus \{0\}$, the space $M_n = \ker(\lambda_n I - K) = \ker(\lambda_n I - K)^2$ is finite dimensional. By Proposition 5.4.4(5), $M_m \perp M_n$ for $1 \leq m < n$.

Let $N = \sum_{n\geq 1}^{\oplus} M_n$ be the orthogonal direct sum of M_n for $n \geq 1$. Then this subspace is invariant for K, and by Proposition 5.4.4(3), it is also invariant for K^*. By Proposition 5.4.4(4), $N^\perp = \bigcap_{n\geq 1} \overline{\operatorname{Ran}(\lambda_n I - K)}$, which is also invariant for K and K^*. So $H = N \oplus N^\perp$ and $K \simeq \begin{bmatrix} K|_N & 0 \\ 0 & K|_{N^\perp} \end{bmatrix}$ because both N and $N^\perp$ are invariant. It is easy to check that $K^* \simeq \begin{bmatrix} K^*|_N & 0 \\ 0 & K^*|_{N^\perp} \end{bmatrix}$. Computation of $0 = K^*K - KK^*$ shows that $K|_N$ and $K|_{N^\perp}$ are both normal. All of the non-zero eigenvectors of K lie in N, so that $\sigma(K|_{N^\perp}) = \{0\}$. By Proposition 5.4.4(2), $\|K|_{N^\perp}\| = 0$. So $K|_{N^\perp} = 0$, which shows that $N^\perp = \ker K = M_0$ and thus $H = \sum_{n\geq 0}^{\oplus} M_n$.

Let P_{M_n} be the orthogonal projection onto M_n. Then $L = \sum_{n\geq 0} \lambda_n P_{M_n}$ is a normal operator with $\ker(\lambda_n I - L) = M_n$ for $n \geq 0$. These spaces sum to the whole space, so L and K agree on each M_n, and therefore on all of H. If we choose an orthonormal basis for each M_n, we obtain an orthonormal basis for H consisting of eigenvalues for K. So K is diagonalized by this basis. ■

Let $K \in \mathcal{K}(H)$ where H is a separable Hilbert space. Then K^*K is a positive compact operator. Hence it is diagonalizable. Thus there is an orthonormal basis, say $\{e_n : n \geq 1\}$ so that $K^*Ke_n = s_n^2 e_n$, where $s_n \geq 0$ and $\lim_{n\to\infty} s_n = 0$. Define a positive compact operator by $|K|e_n = s_n e_n$ for $n \geq 1$. Then $|K|^2 = K^*K$. For $x \in H$,

$$\begin{aligned} \| \, |K|x\|^2 &= \langle |K|x, |K|x\rangle = \langle |K|^2 x, x\rangle \\ &= \langle K^*Kx, x\rangle = \langle Kx, Kx\rangle = \|Kx\|^2. \end{aligned}$$

Define $U_0 : \operatorname{Ran}|K| \to \operatorname{Ran} K$ by $U_0(|K|x) = Kx$. This map is isometric, and so it extends to an isometry from $\overline{\operatorname{Ran}|K|}$ onto $\overline{\operatorname{Ran} K}$. Now $(\operatorname{Ran}|K|)^\perp = \ker|K| = \ker K$; so $H = \overline{\operatorname{Ran}|K|} \oplus \ker K$. Define $U \in \mathcal{B}(H)$ by

$$U(x \oplus y) = U_0 x \quad \text{for} \quad x \in \overline{\operatorname{Ran}|K|} \text{ and } y \in \ker K.$$

Then by construction, $K = U|K|$, and U has the property that $\ker U = \ker K$ and $\operatorname{Ran} U = \overline{\operatorname{Ran} K}$. Algebraically, we have U^*U is the orthogonal projection onto $\overline{\operatorname{Ran}|K|} = (\ker K)^\perp$ and UU^* is the orthogonal projection onto $\overline{\operatorname{Ran} K}$. An operator like U which is an isometry on the orthogonal complement of its kernel is called a *partial isometry*.

We think of the factorization $K = U|K|$ as the *polar decomposition* by analogy with factoring a complex number as $z = e^{i\theta}|z|$. Let $S = \{n : s_n > 0\}$. Then

$$\ker K = \ker|K| = \overline{\operatorname{span}\{e_n : n \notin S\}}$$

and

$$(\ker K)^\perp = \overline{\operatorname{Ran}|K|} = \overline{\operatorname{span}\{e_n : n \in S\}}.$$

Note that $f_n = Ue_n$ for $n \in S$ is an orthonormal set because U_0 is an isometry. We can write $|K| = \sum_{n\in S} s_n e_n e_n^*$. Therefore

$$K = U|K| = U \sum_{n\in S} s_n e_n e_n^* = \sum_{n\in S} s_n f_n e_n^*.$$

The values $(s_n)_{n \in S}$, reordered so that they form a decreasing sequence, are the *singular values* of K. If K is finite rank, then we set $s_k = 0$ for $k > \operatorname{rank} K$.

When H is not separable, a compact operator is always supported on a separable subspace $M = \overline{\operatorname{Ran} K + (\ker K)^\perp}$. Thus K also has a polar decomposition, and the non-zero singular values form a finite or countable set with limit 0.

5.4.7. DEFINITION. The *Schatten p-class* on H is the set

$$\mathfrak{S}_p = \{K \in \mathcal{K}(H) : (s_n) \in l_p\}$$

with norm $\|K\|_p := \|(s_n)\|_p$.

It can be shown that each $\mathfrak{S}_p$ is an ideal of $\mathcal{B}(H)$, and

$$\|SKT\|_p \le \|S\| \, \|K\|_p \, \|T\| \quad \text{for} \quad K \in \mathfrak{S}_p \text{ and } S, T \in \mathcal{B}(H).$$

It contains all finite rank operators, and thus is operator norm dense in $\mathcal{K}(H)$; and it is complete in the $\|\cdot\|_p$ norm. For $1 < p < \infty$, $\mathfrak{S}_p^* = \mathfrak{S}_q$ where $\frac{1}{p} + \frac{1}{q} = 1$. And it is an important theorem that $\mathcal{K}(H)^* = \mathfrak{S}_1$ and $\mathfrak{S}_1^* = \mathcal{B}(H)$. In particular, $\mathcal{B}(H)$ is a dual space, and so has a weak-$*$ topology.

5.5. Invariant Subspaces

5.5.1. DEFINITION. If $\mathcal{A} \subset \mathcal{B}(X)$, a closed subspace $M \subset X$ is an *invariant subspace* for $\mathcal{A}$ if $AM \subset M$ for all $A \in \mathcal{A}$. It is a *proper invariant subspace* if $\{0\} \neq M \neq X$. We write $\operatorname{Lat} \mathcal{A} = \{M : \text{ invariant subspaces for } \mathcal{A}\}$ for the *lattice of invariant subspaces*.

$\operatorname{Lat} \mathcal{A}$ is a *complete lattice*. The operations are intersection, $M \wedge N = M \cap N$, and closed span, $M \vee N = \overline{M + N}$. It is easy to check that these are both invariant subspaces. Completeness indicates that we can take the supremum and infimum over arbitrary subsets of $\operatorname{Lat} \mathcal{A}$. we have

$$\bigwedge_{\alpha \in A} M_\alpha = \bigcap_{\alpha \in A} M_\alpha \quad \text{and} \quad \bigvee_{\alpha \in A} M_\alpha = \overline{\sum_{\alpha \in A} M_\alpha}$$

where in the sum, we allow only finitely many non-zero vectors. Again it is easy to check that these subspaces are invariant.

The following result has a slick proof due to Hilden. The original result is actually stronger.

5.5.2. LOMONOSOV'S THEOREM. *If* $0 \neq K \in \mathcal{K}(X)$*, then* $\{K\}'$ *has a proper invariant subspace.*

PROOF. If $0 \neq \lambda \in \sigma(K)$, then $M = \ker(\lambda I - K)$ is a non-zero finite dimensional (closed) subspace. If $A \in \{K\}'$, then for $x \in M$,

$$(\lambda I - K)Ax = A(\lambda I - K)x = 0.$$

Thus $AM \subset M$; whence $M \in \mathrm{Lat}\{K\}'$.

Otherwise $\sigma(K) = \{0\}$; so that $0 = \mathrm{spr}(K) = \lim_{n\to\infty} \|K^n\|^{1/n}$. We may normalize so that $\|K\| = 1$. Pick a vector $x_0 \in X$ so that $\|Kx_0\| = 1 + \delta > 1$; so $\|x_0\| > 1$. Let $S = b_1(x_0)$, and note that $0 \notin \overline{S}$. Define $D = \overline{KS}$, which is a compact set. If $x \in S$,

$$\|Kx\| \geq \|Kx_0\| - \|K(x - x_0)\| \geq \|Kx_0\| - 1 = \delta > 0.$$

Therefore $0 \notin D$.

Assume that $\{K\}'$ has no proper invariant subspace. For any non-zero vector x, $\overline{\{K\}'x}$ is invariant for $\{K\}'$ and contains x because $\{K\}'$ is an algebra containing the identity. Thus we have that $\overline{\{K\}'x} = X$. Hence for $x \in D$, there is some $A \in \{K\}'$ so that $Ax \in S$. Then $U_A = A^{-1}(S)$ is an open neighbourhood of x. So $\{U_A : A \in \{K\}'\}$ is an open cover of D. Let $U_{A_1}, \dots, U_{A_n}$ be a finite subcover.; i.e., for each $x \in D$, there is some $i \leq n$ so that $A_i x \in S$.

Let $M = \max\{\|A_i\| : 1 \leq i \leq n\}$. Recursively select a sequence $i_1, i_2, i_3, \dots$ in $\{1, 2, \dots, n\}$ so that

$$A_{i_k} K A_{i_{k-1}} K \dots A_{i_1} K x_0 \in S.$$

Therefore since each $A_i \in \{K\}'$,

$$0 < \|x_0\| - 1 \leq \|A_{i_k} \dots A_{i_1} K^k x_0\| \leq M^k \|K^k\| \, \|x_0\|.$$

Take the kth root and take a limit to get

$$1 = \lim_{k\to\infty} (\|x_0\| - 1)^{1/k} \leq \lim_{k\to\infty} M \|K^k\|^{1/k} \, \|x_0\|^{1/k} = 0.$$

This contradiction shows that proper invariant subspaces must exists. ■

5.5.3. EXAMPLE. Consider the Volterra operatos $V \in \mathcal{B}(L^2(0,1))$ from Example 5.1.5, $Vh(x) = \displaystyle\int_0^x h(t)\,dt$. This is compact and has no eigenvlaues. It has the invariant subspace $N_t = \{h \in L^2(0,1) : f|_{(0,t)} = 0 \text{ a.e.}\}$ for $0 \leq t \leq 1$. In fact, these are the only invariant subspaces for V.

5.5.4. COROLLARY. *If $K \in \mathcal{K}(X)$, then there is a maximal chain of subspaces (with respect to containment) which are invariant for $\{K\}'$.*

PROOF. By Zorn's Lemma, there is a maximal chain $\mathcal{N}$ of subspaces in $\mathrm{Lat}\{K\}'$. If $M_\alpha \in \mathcal{N}$ for $\alpha \in A$, then $M = \bigwedge M_\alpha \in \mathrm{Lat}\{K\}'$. Any other subspace $N \in \mathcal{N}$ is either larger than some M_α, and hence $N \supset M$, or is contained in all of them, and hence $N \subset M$. In either case, it shows that $\mathcal{N} \cup \{M\}$ is a chain of subspaces;

and thus $M \in \mathcal{N}$ by maximality. Similarly, $\bigvee M_\alpha \in \mathcal{N}$. Thus the chain $\mathcal{N}$ is complete.

If $N \in \mathcal{N}$, let $N_- = \bigvee\{M \in \mathcal{N} : M \subsetneq N\}$. If $N_- \subsetneq N$, I claim that $\dim N/N_- = 1$. Otherwise, we define an operator

$$\tilde{K} \in \mathcal{B}(N/N_-) \quad \text{by} \quad \tilde{K}(x + N_-) = Kx + N_-.$$

Note that $\overline{\tilde{K}b_1(N/N_-)} = Q(\overline{Kb_1(N)})$, where $Q : N \to N/N_-$ is the quotient map. The continuous image of a compact set is compact, and therefore $\tilde{K}$ is a compact operator. The algebra $\{\tilde{K}\}'$ has a proper invariant subspace, say M. Then $Q^{-1}M =: L$ is invariant for $\{K\}'$ and $N_- \subsetneq L \subsetneq N$. The reason is that for $T \in \{K\}'$, the operator $\tilde{T}(x + N_-) = Tx + N_-$ in $\mathcal{B}(N/N_-)$ lies in $\{\tilde{K}\}'$ and so leaves M invariant. Pulling this back to X shows that L is invariant for T. However this contradicts the maximality of $\mathcal{N}$. Hence the gap N/N_- is either 0 or 1 dimensional.

This chain $\mathcal{N}$ must be a maximal chain in the lattice of all subspaces of X. For if $\mathcal{N}\dot\cup\{M\}$ were a larger chain of subspaces, let $N_- = \bigvee\{N \in \mathcal{N} : N \subset M\}$ and $N_+ = \bigwedge\{N \in \mathcal{N} : M \subset N\}$. Then by completeness, $N_\pm \in \mathcal{N}$ and $N_- \subsetneq M \subsetneq N_+$. Moreover $N_- = (N_+)_-$ and the gap N_+/N_- is at least 2-dimensional. This is false, and hence $\mathcal{N}$ is maximal. ∎

For the rest of this section, we will restrict our attention to operators on Hilbert space. This is for convenience of presentation.

5.5.5. Definition. A *nest of subspaces* is a complete chain of subspaces. The *nest algebra* $\mathcal{T}(\mathcal{N}) = \{T \in \mathcal{B}(H) : TN \subset N \text{ for all } N \in \mathcal{N}\}$. Again we define $N_- = \bigvee\{M \in \mathcal{N} : M \subsetneq N\}$. The *atoms* of $\mathcal{N}$ are the subspaces $A = N \ominus N_- = \{x \in N : x \perp N_-\}$ when $N_- \subsetneq N$; and let $\mathbb{A}$ denote the set of all atoms of $\mathcal{N}$. Define $\Phi_A(T) = P_A T|_A$ be the compression of $T \in \mathcal{T}(\mathcal{N})$ to the subspace A. When $\mathcal{N}$ is a maximal nest, the atoms are 1-dimensional and so $\Phi_A(T)$ are scalars.

Note that $\mathcal{T}(\mathcal{N})$ is a vector space closed under multiplication and contains I. If $T_\alpha \in \mathcal{T}(\mathcal{N})$ converge in the weak operator topology to T, i.e. $\langle T_\alpha x, y\rangle \to \langle Tx, y\rangle$ for each $x, y \in H$, then for every $x \in N$ and $y \in N^\perp$, we have $0 = \langle T_\alpha x, y\rangle$, and thus $0 = \langle Tx, y\rangle$. Therefore $Tx \in (N^\perp)^\perp = N$; so $T \in \mathcal{T}(\mathcal{N})$. Therefore $\mathcal{T}(\mathcal{N})$ is a weak operator topology closed algebra of operators. The fact that each $N \in \mathcal{N}$ is invariant can be seen as an upper triangular form for elements of $\mathcal{T}(\mathcal{N})$.

We will use the notation of Theorem 5.2.7. The proof will be given in a series of lemmas.

5.5.6. RINGROSE'S THEOREM. *Suppose that $K \in \mathcal{K}(H)$ and that $\mathcal{N}$ is a maximal nest in* Lat K*. Let $\mathbb{A}$ be the atoms of $\mathcal{N}$. Then*

$$\sigma(K) = \{0\} \cup \{\Phi_A(K) : A \in \mathbb{A}\}.$$

Moreover each non-zero eigenvalue λ is repeated $n_\lambda = \dim N(\lambda)$ times.

5.5.7. LEMMA. $\sigma(K) = \{0\} \cup \{\Phi_A(K) : A \in \mathbb{A}\}$.

PROOF. Let $A \in \mathbb{A}$ such that $0 \neq \lambda = \Phi_A(K)$. Write $A = N \ominus N_-$ for some $N \in \mathcal{N}$. Since A is 1-dimensional, pick a unit vector $x \in A$. Then $\lambda = \Phi_A(K) = \langle Kx, x\rangle$. Since $N = N_- \oplus \mathbb{C}x$, $Kx = \lambda x + y$ for some $y \in N_-$. Hence $(\lambda I - K)x = -y \in N_-$. This shows that $(\lambda I - K)|_N$ is not surjective, and so $\lambda \in \sigma(K|_N)$. By Theorem 5.2.7, $\lambda \in \sigma_p(K|_N)$. Thus $\ker(\lambda I - K) \supset \ker(\lambda I - K)|_N \neq \{0\}$. Hence $\lambda \in \sigma(K)$.

Conversely suppose that $\lambda \in \sigma(K)\setminus\{0\}$. Then $E = \ker(\lambda I - K)$ is a non-zero finite dimensional subspace. Thus the sphere $S = \{x \in E : \|x\| = 1\}$ is compact. Let

$$N = \bigwedge \{M \in \mathcal{N} : E \cap M \neq \{0\}\}.$$

If $E \cap M \neq \{0\}$, then $S \cap M$ is non-zero and compact. This collection has the finite intersection property because $(S \cap M_1) \cap (S \cap M_2) = S \cap (M_1 \cap M_2)$, and $M_1 \cap M_2$ is the smaller of M_1 and M_2. By compactness, the intersection $S \cap N$ is non-empty. Thus $E \cap N \neq \{0\}$. Therefore $\lambda I - K|_N$ has non-trivial kernel, and thus $\operatorname{Ran}(\lambda I - K|_N)$ is a proper closed subspace of N.

Now if $M \in \mathcal{N}$ and $M < N$, then $E \cap M = \{0\}$. Hence $\lambda I - K|_M$ is injective, and thus by Theorem 5.2.7, $\operatorname{Ran}(\lambda I - K|_M) = M$. Since N_- is the closed span of $M \in \mathcal{N}$ with $M < N$, $\operatorname{Ran}(\lambda I - K|_{N_-}) \supset \bigcup_{M<N} M$. This is dense in N_- and the range is closed, so equals N_-. This shows that $N_- \subsetneq N$. It follows that $\operatorname{Ran}(\lambda I - K|_N) \supset N_-$ and is a proper subspace of N. Since $A = N \ominus N_-$ is 1-dimensional, the range is N_-. Therefore $0 = \Phi_A(\lambda I - K) = \lambda - \Phi_A(K)$. Thus $\Phi_A(K) = \lambda$. ∎

5.5.8. LEMMA. *Let $\lambda \in \sigma(K) \setminus \{0\}$, and let $M \in$ Lat K. Then M is invariant for E_λ.*

PROOF. The restriction $K|_M$ is also compact. Let

$$N = \ker(\lambda I - K|_M)^{n_\lambda} = N(\lambda) \cap M \quad \text{and} \quad R = \operatorname{Ran}(\lambda I - K|_M)^{n_\lambda} \subset R(\lambda) \cap M.$$

Then by Theorem 5.2.7, $M = N \dotplus R$. Note that $E_\lambda|_N = I|_N$ and $E_\lambda|_R = 0$. Therefore $E_\lambda M = N \subset M$. ∎

5.5.9. LEMMA. *Let $\lambda \in \sigma(K) \setminus \{0\}$, and let $A \in \mathbb{A}$. Then*

$$\Phi_A(E_\lambda) = \begin{cases} 1 & \text{if } \Phi_A(K) = \lambda \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. We have $H = N(\lambda) \dotplus R(\lambda)$ is a direct sum of subspaces in $\operatorname{Lat} K$. Then $K_\lambda := K|_{N(\lambda)}$ has the form $\lambda I_\lambda + J$, where I_λ is the identity on $N(\lambda)$ and $J \in \mathcal{B}(N(\lambda))$ is nilpotent. Let $L = K|_{R(\lambda)}$. This is compact and $\lambda \notin \sigma(L)$. With respect to the decomposition $H = N(\lambda) \dotplus R(\lambda)$, K, E_λ and $K + zE_\lambda$ have matrices

$$K = \begin{bmatrix} \lambda I_\lambda + J & 0 \\ 0 & L \end{bmatrix}, \quad E = \begin{bmatrix} I_\lambda & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad K + tE_\lambda = \begin{bmatrix} (\lambda + z)I_\lambda + J & 0 \\ 0 & L \end{bmatrix}.$$

By Lemma 4.4.8,

$$\sigma(K + zE_\lambda) = \sigma((\lambda + z)I_\lambda + J) \cup \sigma(L) = \{\lambda + z\} \cup \sigma(K) \setminus \{\lambda\}.$$

For any $A \in \mathbb{A}$ and $t \in \mathbb{C}$, we have that

$$\Phi_A(K + zE_\lambda) = \Phi_A(K) + z\Phi_A(E_\lambda) \in \{\lambda + z\} \cup \sigma(K) \setminus \{\lambda\}.$$

If $\Phi_A(K) = \lambda$, then $\Phi_A(K + zE_\lambda)$ will be near to λ for small z, and therefore $\Phi_A(K+zE_\lambda) = \lambda+z$. Thus $\Phi_A(E_\lambda) = 1$. On the other hand, if $\Phi_A(K) = \mu \neq \lambda$, then $\mu + z\Phi_A(E_\lambda) \in \sigma(K) \setminus \{\lambda\}$ for very large z, and hence $\Phi_A(E_\lambda) = 0$. ■

5.5.10. LEMMA. *The set $\mathbb{A}(\lambda) = \{A \in \mathbb{A} : \Phi_A(K) = \lambda\}$ has cardinality $n_\lambda = \operatorname{rank} E_\lambda$.*

PROOF. Let $\mathbb{A}(\lambda) = \{A_1, \dots, A_n\}$ be a finite or countable set, and define $P = \sum_{A_i \in \mathbb{A}(\lambda)} A_i$. This is an orthogonal projection of rank $|\mathbb{A}(\lambda)|$. Each $A_i = N_i \ominus N_{i-} = \mathbb{C}e_i$. Then $E_\lambda e_i = e_i + y_i$ where $y_i \in N_{i-}$. In particular, if we order the N_i by containment, then we see that the compression $PE_\lambda|_{PH}$ is upper triangular with 1's on the diagonal. Therefore $n \leq \operatorname{rank} E_\lambda$.

The projection P also lies in $\mathcal{T}(\mathcal{N})$ since either $e_i \in N$ or $e_i \in N^\perp$. By choice of P, we have $\Phi_A(E_\lambda - P) = 0$ for $A \in \mathbb{A}(\lambda)$. Also by construction, $\Phi_A(P) = 0 = \Phi_A(E_\lambda)$ for $A \in \mathbb{A} \setminus \mathbb{A}(\lambda)$. By Lemma 5.5.7 applied to the finite rank operator $E_\lambda - P$, $\sigma(E_\lambda - P) = \{0\}$. If $\operatorname{rank} P < \operatorname{rank} E_\lambda$, there would be a unit vector x in $N(\lambda) \cap P^\perp H$, and $(E_\lambda - P)x = x$. Hence $|\mathbb{A}(\lambda)| = \operatorname{rank} P = \operatorname{rank} E_\lambda = n_\lambda$. ■

For vectors $x, y \in H$, we use the notation xy^* for the rank one operator $(xy^*)(z) = \langle z, y\rangle x$. This is a natural notation because

$$xy^* = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{bmatrix} \begin{bmatrix} \overline{y_1} & \overline{y_2} & \overline{y_3} & \dots \end{bmatrix} = \begin{bmatrix} x_1\overline{y_1} & x_1\overline{y_2} & x_1\overline{y_3} & \dots \\ x_2\overline{y_1} & x_2\overline{y_2} & x_2\overline{y_3} & \dots \\ x_3\overline{y_1} & x_3\overline{y_2} & x_3\overline{y_3} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Also we have the identity $A(xy^*)B = (Ax)(B^*y)^*$ for any $A, B \in \mathcal{B}(H)$.

5.5.11. LEMMA. *Let* $\mathcal{A} \subset \mathcal{B}(H)$. *Then* $\operatorname{Lat}\mathcal{A}^* = \{M^\perp : M \in \operatorname{Lat}\mathcal{A}\}$.

PROOF. If $M \in \operatorname{Lat}\mathcal{A}$, $x \in M$, $y \in M^\perp$ and $A \in \mathcal{A}$, then $\langle A^*y, x\rangle = \langle y, Ax\rangle = 0$. This shows that $\mathcal{A}^*M^\perp$ is orthogonal to M, whence $\mathcal{A}^*M^\perp \subset M^\perp$. Since $\mathcal{A} = (\mathcal{A}^*)^*$, this argument is reversible. ■

We finish this section with a useful finite dimensional result.

5.5.12. BURNSIDE'S THEOREM. *If* $\mathcal{A} \subset \mathcal{M}_n$ *is a subalgebra of the* $n \times n$ *complex matrices. and* $\mathcal{A}x = \mathbb{C}^n$ *for all* $x \neq 0$, *then* $\mathcal{A} = \mathcal{M}_n$.

PROOF. Proceed by induction on n. For $n = 1$, the only proper subalgebra is $\{0\}$, which doesn't satisfy the hypothesis. So the result is true. Assume the result is true for all $k < n$.

First I show that there is some $F \in \mathcal{A}$ with $0 < \operatorname{rank} F < n$. $\mathcal{A}$ cannot consist solely of scalar multiples of I since $n \geq 2$, so pick a non-scalar $A \in \mathcal{A}$. If A is not invertible, then $0 < \operatorname{rank} A < n$ as desired. Otherwise pick $\lambda \in \sigma(A)$, so $A - \lambda I$ is non-zero and not invertible. The rank is unchanged if we multiply by the invertible A, so $F = A(A - \lambda) = A^2 - \lambda A \in \mathcal{A}$ works.

Let $M = \operatorname{Ran} F$. The set $\mathcal{B} = F\mathcal{A}|_M$ is an algebra since

$$(FA|_M)(FB|_M) = F(AFB)|_M \in \mathcal{B} \quad \text{for} \quad A, B \in \mathcal{A}.$$

If $0 \neq x \in M$, then $\mathcal{B}x = F\mathcal{A}x = F\mathbb{C}^n = M$. By the induction hypothesis, $\mathcal{B} = \mathcal{L}(M)$. In particular, $\mathcal{B}$ contains a rank one operator $R = FA_0|_M$. Therefore $FA_0F \in \mathcal{A}$ has rank 1, say $FA_0F = x_0y_0^*$.

Let $x, y \in \mathbb{C}^n$. Pick $A \in \mathcal{A}$ so that $Ax_0 = x$. Now $\mathcal{A}^*y_0$ is invariant for $\mathcal{A}^*$, so by Lemma 5.5.11, it is $\{0\}$ or $\mathbb{C}^n$. But if $\mathcal{A}^*y_0 = 0$, then

$$0 = \langle A^*y_0, x_0\rangle = \langle y_0, \mathcal{A}x_0\rangle$$

which contradicts that $y_0 \in \mathcal{A}x_0$. Pick $B^* \in \mathcal{A}^*$ so that $B^*y_0 = y$. Then $\mathcal{A}$ contains $A(x_0y_0^*)B = (Ax_0)(B^*y_0)^* = xy^*$. Every matrix is a sum of rank one matrices, so $\mathcal{A} = \mathcal{M}_n$. ■

Exercises for Chapter 5

1. (a) If $E = E^2 \in \mathcal{B}(X)$ is compact, show that E is finite rank.
(b) Show that if $f \in C[0,1]$, then the multiplication operator M_f on $L^2(0,1)$ is not compact unless $f = 0$.

2. (a) Let X, Y be Banach spaces. Prove that if $K \in \mathcal{K}(X,Y)$ is compact, then whenever a *sequence* $x_n \in X$ converges weakly to x_0, the sequence Kx_n converges to Kx_0 in norm.
(b) Let X be a *separable reflexive* Banach space. Suppose $T \in \mathcal{B}(X,Y)$ and whenever a sequence $x_n \in X$ converges weakly to x_0, the sequence Tx_n converges to Tx_0 in norm. Prove that T is compact. **Hint:** use the Banach–Alaoglu Theorem.
(c) Find a *non-compact* operator $T \in \mathcal{B}(l_2)$ so that $\lim\limits_{n\to\infty} \|Te_n\| = 0$ for the usual o.n. basis $\{e_n\}$.

3. Let V be a bounded operator on a Hilbert space H.
(a) Show that V is an isometry if and only if $V^*V = I$.
(b) Show that the following are equivalent:
(i) U is unitary, (ii) $U^* = U^{-1}$, and (iii) U is an isometry and $UU^* = U^*U$.

4. Suppose that $T \in \mathcal{B}(X,Y)$ is Fredholm. Prove that T^* is Fredholm, and find the relationship between $\operatorname{ind} T$ and $\operatorname{ind} T^*$.

5. Let S be the unilateral shift on l_2. Define $T = S \oplus (S^* + \frac{1}{2}I)$ in $\mathcal{B}(l_2 \oplus l_2)$. Compute $\operatorname{ind}(T - \lambda I)$ for $\lambda \in \{-\frac{3}{4}, 0, \frac{5}{4}\}$.

6. The Donaghue operator acts on l_2 by $Ae_0 = 0$ and $Ae_n = 2^{-n}e_{n-1}$ for all $n \geq 1$.
(a) Show that A is compact, and compute $\|A\|$.
(b) Show that the proper closed invariant subspaces for A are precisely the subspaces $\mathcal{M}_n = \operatorname{span}\{e_k : 0 \leq k \leq n\}$ for $n \geq 0$. HINT: If an invariant subspace $\mathcal{M}$ contains a vector $x = \sum_{n\geq 0} x_n e_n$ with $x_n \neq 0$ infinitely often, show that $e_0 \in \mathcal{M}$. Pick $n_i \to \infty$ such that $|x_{n_i}| \geq |x_n|$ for all $n \geq n_i$. Consider appropriate multiples of $A^{n_i}x$.

7. Let V be the Volterra operator on $L^2(0,1)$, $Vf(x) = \int_0^x f(t)\,dt$.
(a) Express V^* as an integral operator.
(b) Suppose that f is an eigenvector of VV^*. Show that f is C^∞ and satisfies a second order ODE with boundary conditions.
(c) Hence diagonalize VV^* and compute $\|V\|$.

Index