Notes for Math 371 (Undergraduate Functional Analysis)
©Department of Mathematics, The Hong Kong University of Science and Technology

References

These notes are based on materials in the following books:


Also, we will cite some results from the books below:


Abbreviations and Notations

iff if and only if
end of proof
\( \mathbb{R} \) or \( \mathbb{C} \)
Chapter 0. Set and Topological Preliminaries.

§1. Axiom of Choice and Zorn’s Lemma. We begin by introducing the following axiom from set theory.

**Axiom of Choice.** Let $A$ be a nonempty set and for every $\alpha \in A$, let $S_\alpha$ be a nonempty set. For $S = \{S_\alpha : \alpha \in A\}$, there exists a function $f : A \to \bigcup S = \bigcup\{S_\alpha : \alpha \in A\}$ such that for all $\alpha \in A$, $f(\alpha) \in S_\alpha$.

From this we can deduce Zorn’s lemma, which is a powerful tool in showing the existence of many important objects. To set it up, we need some terminologies.

**Definitions.** (1) A relation $R$ on a set $X$ is a subset of $X \times X$.

(2) For a relation $R$, we now write $x \leq y$ (or $y \geq x$) iff $(x, y) \in R$. Also, $x < y$ iff $x \leq y$ and $x \neq y$. $R$ is a partial ordering of $X$ if it satisfies the reflexive property ($x \leq x$ for all $x \in R$), the antisymmetric property ($x \leq y$ and $y \leq x$ imply $x = y$) and the transitive property ($x \leq y$ and $y \leq z$ imply $x \leq z$). $X$ is a poset (or a partially ordered set) iff there is a partial ordering $R$ on $X$.

(3) A poset $X$ is totally ordered (or linearly ordered or simply ordered) iff for all $x, y \in X$, either $x \leq y$ or $y \leq x$.

(4) A poset $X$ is well-ordered iff every nonempty subset $G$ of $X$ has a least element in $G$, i.e. there is $g_0 \in G$ such that for all $g \in G$, $g_0 \leq g$. (Taking $G = \{x, y\}$, we see $X$ well-ordered implies $X$ totally ordered.)

(5) A chain in a poset $X$ is either the empty set or a totally ordered subset of $X$.

(6) An element $u$ in a poset $X$ is an upper bound for a subset $S$ of $X$ iff $x \leq u$. An element $m$ of $X$ is maximal in $X$ iff $m \leq x$ implies $x = m$. (Similarly lower bound and minimal element may be defined.)

**Examples.** (1) For $X = \mathbb{R}$ with the usual ordering (i.e. $x \leq y$ iff $x \leq y$), $\mathbb{R}$ is totally ordered. $(0, \infty)$ is a chain in $\mathbb{R}$ with no upper bound in $\mathbb{R}$. $\mathbb{R}$ has no maximal element.

(2) For every set $W$, the power set $X = P(W) = \{S : S \subseteq W\}$ has a partial ordering given by inclusion (i.e. $S \subseteq T$ iff $S \subseteq T$). It is not totally ordered when $W$ has more than one elements. For distinct elements $a, b$ of $W$, then neither $\{a\} \leq \{b\}$ nor $\{b\} \leq \{a\}$. $W$ is the unique maximal element in $X = P(W)$.

(3) For the closed unit disc $X = D = \{(x, y) : x^2 + y^2 \leq 1\}$ of $\mathbb{R}^2$, define $(x_1, y_1) \leq (x_2, y_2)$ if $x_1 \leq x_2$ and $y_1 \leq y_2$. Every point on the minor arc of the unit circle joining $(1, 0)$ to $(0, 1)$ is a maximal element of $D$.

**Zorn’s Lemma.** For a nonempty poset $X$, if every chain in $X$ has an upper bound in $X$, then $X$ has at least one maximal element. (The statement is also true if ‘upper’ and ‘maximal’ are replaced by ‘lower’ and ‘minimal’ respectively.)

For a proof, see the appendix at the end of the chapter. Below we will present two examples of Zorn’s lemma, namely (1) for any two nonempty sets, there exists an injection from one of them to the other and (2) every nonzero vector space has a basis.

**Remark.** If $X$ is a nonempty collection of subsets of some set $W$ and we define the set inclusion relation $R = \{(A, B) \mid A, B \in X, A \subseteq B\}$ on $X$ (i.e. $A \subseteq B$ if $A \subseteq B$), then we can check $X$ is partially ordered by this relation:

(a) For every $A \in X$, we have $A = A \implies A \subseteq A$.

(b) For every $A, B \in X$, we have $A \subseteq B$ and $B \subseteq A \implies A = B$.

(c) For every $A, B, C \in X$, we have $A \subseteq B$ and $B \subseteq C \implies A \subseteq C$.

**Example 1.** For nonempty sets $A$ and $B$, there exists an injective function either from $A$ to $B$ or from $B$ to $A$.
Proof. Let $W = A \times B$. Let $g : C \to B$ be a function, where $C \subseteq A$, then $\Gamma(g) = \{ (c, g(c)) \mid c \in C \} \subseteq W$. Let $X = \{ \Gamma(g) \mid g : C \to B \text{ is injective, where } \emptyset \subset C \subseteq A \}$. Define the set inclusion relation on $X$, i.e. $\Gamma(g_0) \subseteq \Gamma(g_1)$ if $\Gamma(g_0) \subseteq \Gamma(g_1)$. By the remark above, this is a partial ordering on $X$.

Next for every chain $C = \{ \Gamma(g_{\alpha}) \mid \alpha \in I, \ g_{\alpha} : C_{\alpha} \to B \text{ is injective, where } \emptyset \subset C_{\alpha} \subseteq A \}$ in $X$, we will show $S = \bigcup_{\alpha \in I} \Gamma(g_{\alpha})$ is in $X$. (Observe that a nonempty subset $T$ of $W = A \times B$ is an element of $X$ iff for every pair of distinct points $(a, b), (a', b')$ in $T$, we have $a \neq a'$ (by the definition of function) and $b \neq b'$ (by injectivity).)

Let $(a, b)$ and $(a', b')$ be distinct points in $S$. Then there are $\alpha, \alpha' \in I$ such that $(a, b) \in \Gamma(g_{\alpha})$ and $(a', b') \in \Gamma(g_{\alpha'})$. Since $C$ is a chain in $X$, we may suppose $\Gamma(g_{\alpha}) \subseteq \Gamma(g_{\alpha'})$. Then $(a, b)$ and $(a', b')$ are distinct points in $\Gamma(g_{\alpha'})$. Since $g_{\alpha}$ is injective, $a \neq a'$ and $b \neq b'$. Therefore, $S$ is in $X$. Finally, since for all $\alpha \in I$, $\Gamma(g_{\alpha}) \subseteq S$, so $S$ is an upper bound of $C$.

By Zorn’s lemma, $X$ has a maximal element $M = \Gamma(f)$. We claim that either the domain of $f$ is $A$ or the range of $f$ is $B$. Assume not, then there exist $a \in A$ not in the domain of $f$ and $b \in B$ not in the range of $f$. It follows $M' = M \cup \{(a, b)\}$ is in $X$ and $M \subseteq M'$, a contradiction. So the claim is true.

If the domain of $f$ is $A$, then $f : A \to B$ is injective. If the range of $f$ is $B$, then $f^{-1} : B \to A$ is injective.

Example 2. Every nonzero vector space $W$ over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ has a basis.

Proof. For a subset $S$ of $W$, recall that $S$ is linearly independent iff every finite subset of $S$ is linearly independent. Let $X = \{ S \mid S \text{ is a linearly independent subset of } W \}$. By the remark above, the set inclusion relation on $X$ is a partial ordering.

For every chain $C = \{ S_{\alpha} \mid \alpha \in I \}$ in $X$, let $S_I = \bigcup_{\alpha \in I} S_{\alpha}$. We will check $S_I$ is in $X$. For every finite subset $\{ x_1, x_2, \ldots, x_n \}$ in $S_I$, there are $\alpha_1, \alpha_2, \ldots, \alpha_n \in I$ such that $x_1 \in S_{\alpha_1}, x_2 \in S_{\alpha_2}, \ldots, x_n \in S_{\alpha_n}$. Since $C$ is a chain, we may assume $S_{\alpha_2}, \ldots, S_{\alpha_n} \subseteq S_{\alpha_1}$. Then $\{ x_1, x_2, \ldots, x_n \}$ in $S_{\alpha_1}$. Since $S_{\alpha_1}$ is linearly independent, so $\{ x_1, x_2, \ldots, x_n \}$ is linearly independent. Therefore, $S_I$ is in $X$. Clearly, $S_I$ is an upper bound of $C$.

By Zorn’s lemma, $X$ has a maximal element $M$. We claim that the span of $M$ is $W$. Assume there exists $x \in W$ not in the span of $M$. By the maximality of $M$, $M' = M \cup \{ x \}$ cannot be in $X$, i.e. $M'$ is not linearly independent. So there exists $x_1, x_2, \ldots, x_n \in M$ and $c_1, c_2, \ldots, c_n, c \in \mathbb{K}$ (not all zeros) such that $c_1x_1 + c_2x_2 + \cdots + c_nx_n + cx = 0$. Since $M$ is linearly independent, we must have $c \neq 0$. Then $x = \frac{1}{c}(c_1x_1 + c_2x_2 + \cdots + c_nx_n)$ is in the span of $M$, a contradiction. So the claim is true.

Finally, since $M \in X$ is linearly independent and $M$ spans $W$, $M$ is a basis of $W$.

Exercises. (1) Prove that there exists a collection $S$ of pairwise disjoint open disks on a plane such that every open disk on the plane must intersect at least one open disk in $S$. (Hint: Partial order collections consisted of pairwise disjoint open disks.)

(2) Prove that for every integer $n \geq 3$, there exist a set $S_n \subseteq [0, 1]$ such that $S_n$ contains no $n$-term arithmetic progression, but for every $x \in [0, 1] \setminus S_n$, $S_n \cup \{ x \}$ contains a $n$-term arithmetic progression.

(3) Prove that a normed space $X$ is nonseparable if and only if there exists uncountably many pairwise disjoint open balls of radius $1$ in $X$.

Remarks. (1) Actually the axiom of choice and Zorn’s lemma (as well as a few other principles from set theory) are equivalent, see [HS], pp. 14-17.

(2) Zorn’s lemma also holds if antisymmetric property of a partial ordering is omitted. See [M], p. 8, ex. 1.16. If ‘chain’ is replaced by ‘well-ordered subset’ everywhere, Zorn’s lemma and the proof are still correct.

(3) The axiom of choice is used to prove that every set of positive outer Lebesgue measure in $\mathbb{R}$ has non-measurable subsets. (See [Ku], pp. 287-288.) Important applications of Zorn’s lemma include the following:
(a) Every nonzero Hilbert space has an orthonormal basis. (See [RS], pp. 44-45.)
(b) In every nonzero ring with an identity, every ideal is contained in a maximal ideal. (See [H], p. 128.)
(c) Every field has an algebraic closure. (See [Mc], pp. 21-22.)

§2. Topology. In the sequel, the prepositional phrase a set $S$ in $X$ will mean $S \subseteq X$. Now we begin by introducing the concept of topology on a set $X$, which generalizes the concept of all open sets in $\mathbb{R}$.

**Definitions.** (1) Let $X$ be a set and $T$ be a collection of subsets of $X$. $T$ is a topology on $X$ iff

(a) $\emptyset, X \in T$,
(b) the union of any collection of elements of $T$ is an element of $T$,
(c) the intersection of finitely many elements of $T$ is an element of $T$.

A set $X$ with a topology is called a topological space. In case the topology is clear, we simply say $X$ is a topological space.

(2) Let $S \subseteq X$. $S$ is open in $X$ iff $S \in T$. $S$ is closed in $X$ iff $X \setminus S \in T$. (Using de Morgan’s law, we can get topological properties for closed sets, namely (a’) $\emptyset$, $X$ are closed, (b’) the intersection of any collection of closed sets is closed and (c’) the union of finitely many closed sets is closed.)

(3) Let $S \subseteq X$. The interior $S^0$ of $S$ is the union of all open subsets of $S$. (This is the largest open subset of $S$.) The closure $\overline{S}$ of $S$ is the intersection of all closed sets containing $S$. (This is the smallest closed set containing $S$.) $S$ is dense iff $\overline{S} = X$ (equivalently every nonempty open set in $X$ contains a point of $S$).

(4) For every $x \in X$, a subset $N$ of $X$ is a neighborhood of $x$ iff there exists $U \in T$ such that $x \in U \subseteq N$.

(5) A subset $T_0$ of a topology $T$ on $X$ is a base of $T$ iff whenever $x \in U \in T$, there exists $V \in T_0$ such that $x \in V \subseteq U$ (cf Exercise (4) below).

**Remark.** When we are dealing with more than one topologies $T_1, T_2, \ldots$, we shall refer to the elements of $T_1$ as $T_1$-open sets, the elements of $T_2$ as $T_2$-open sets, etc.

**Examples.** (1) If $T_1, T_2$ are topologies on $X$ and $T_1 \subseteq T_2$, then we say $T_1$ is weaker than $T_2$ (or $T_2$ is stronger than $T_1$). For every set $X$, there is a weakest topology on $X$ consisted of $\emptyset$ and $X$. It is called the indiscrete topology on $X$. Also, there is a strongest topology on $X$ consisted of the collection $P(X)$ of all subsets of $X$. This is called the discrete topology on $X$.

(2) The set of all open sets in a metric space $M$ is a topology on $M$. It is called the metric topology on $M$. In the case $M = \mathbb{R}^n$ with the usual metric, it is called the usual topology. The set of all open balls is a base of the metric topology on $M$. Now every open set in $M$ is a union of open balls. This is true in general.

**Exercises.** (4) Prove that a subset $T_0$ of the topology $T$ on $X$ is a base if and only if every open set is a union of elements of $T_0$.

(5) Prove that a collection $B$ of subsets of $X$ is a base of a topology on $X$ if and only if $\bigcup_{V \in B} V = X$ and for every $V_0, V_1 \in B$ and $x \in V_0 \cap V_1$, there exists $V_2 \in B$ such that $x \in V_2 \subseteq V_0 \cap V_1$. (See [D], pp. 47-48.)

§§2.1. Compactness. We now introduce a main concept, namely compactness, in analysis.

**Definitions.** For $S \subseteq X$, a collection $J$ of open sets is an open cover of $S$ iff the union of the elements of $J$ contains $S$. $S$ is compact in $X$ iff every open cover $J$ of $S$ has a finite subset $J_0$ which is also an open cover of $S$. (Such $J_0$ is a finite subcover of $J$.) $S$ is precompact (or relatively compact) if the closure of $S$ is compact.
Definitions. If $X$ is a topological space with topology $\mathcal{T}$ and $W \subseteq X$, then $\mathcal{T}_W = \{ S \cap W : S \in \mathcal{T} \}$ is a topology on $W$ called the relative topology on $W$. A subset $V$ of $W$ is open in $W$ iff $V \in \mathcal{T}_W$. If $\mathcal{B}$ is a base of $\mathcal{T}$, then $\mathcal{B}_W = \{ S \cap W : S \in \mathcal{B} \}$ is a base of $\mathcal{T}_W$.

Remarks. For $V \subseteq W \subseteq X$, if $V$ is open (or closed) in $X$, then $V = V \cap W$ is open (or closed) in $W$, respectively. The converse is false as $(0, 1)$ is open and closed in $(0, 1]$, but neither open nor closed in $\mathbb{R}$.

Intrinsic Property of Compactness. Let $X$ be a topological space with topology $\mathcal{T}$ and $W \subseteq X$. $W$ is compact in $W$ with the relative topology $\mathcal{T}_W$ iff $W$ is compact in $X$ with topology $\mathcal{T}$.

Proof. A collection $J$ of open sets in $X$ covers $W$ in $X$ iff $J_W = \{ S \cap W : S \in J \}$ covers $W$ in $W$. $J$ has a finite subcover iff $J_W$ has a finite subcover.

Remark. Applying de Morgan’s law, $S$ compact in $X$ (equivalently, in $S$) if and only if every collection $\mathcal{F}$ of closed sets in $S$ having the finite intersection property (i.e. the intersection of finitely many members of $\mathcal{F}$ is always nonempty) must satisfy $\bigcap \{ W : W \in \mathcal{F} \} \neq \emptyset$.

§§2.2. Continuity. Observe that if $a < b$ in $\mathbb{R}$, then $(-\infty, (a+b)/2)$ and $(a+b)/2, +\infty)$ are disjoint open sets separating $a$ and $b$. This is a property that makes limit unique if it exists. So we introduce the following.

Definition. A set $X$ with a topology $\mathcal{T}$ is a Hausdorff space (or a $T_2$-space) iff for every distinct $a, b \in X$, there exist disjoint $U, V \in \mathcal{T}$ such that $a \in U$ and $b \in V$.

Once we have topologies on sets, we can study “continuous” functions between them.

Definitions. Let $X, Y$ be topological spaces with topologies $\mathcal{T}_X, \mathcal{T}_Y$ respectively.

(1) $f : X \to Y$ is continuous at $x$ iff for every neighborhood $N$ of $f(x)$, $f^{-1}(N)$ is a neighborhood of $x$. $f : X \to Y$ is continuous iff for every $\mathcal{T}_Y$-open set $U$ in $Y$, $f^{-1}(U)$ is a $\mathcal{T}_X$-open set in $X$ (equivalently, for every $\mathcal{T}_X$-closed set $V$ in $Y$, $f^{-1}(V)$ is a $\mathcal{T}_X$-closed set in $X$).

(2) $f : X \to Y$ is a homeomorphism iff $f$ is bijective and both $f$ and $f^{-1}$ are continuous. (In this case, $U$ is open in $X$ iff $f(U)$ is open in $Y$. We say $X$ and $Y$ are homeomorphic.)

Exercises. Prove the following elementary properties (see [Be], pp. 15, 34-35).

(6) If $f : X \to Y$ and $g : Y \to Z$ are continuous, then $g \circ f : X \to Z$ is continuous.

(7) If $S$ is compact and $X$ is a closed subset of $S$, then $X$ is compact.

(8) If $S$ is Hausdorff and $Y$ is a compact subset of $S$, then $Y$ is closed.

(9) Let $f : X \to Y$ be continuous. If $X$ is compact, then $f(X)$ is compact.

(10) Let $X$ be compact and $Y$ be Hausdorff. If $f : X \to Y$ is continuous and bijective, then $f$ is a homeomorphism.

§§2.3. Nets and Convergence. In metric space, we know that the closure of a set is consisted of all limits of sequences in the set. However, this is false in general for topological spaces as shown by the following example!

Example. On $[0, 1]$, define open sets to be either empty or sets whose complements in $[0, 1]$ are countable. More precisely, let $\mathcal{T} = \{ \emptyset \} \cup \{ S : S \subseteq [0, 1], [0, 1] \setminus S \text{ is countable} \}$. We can check $\mathcal{T}$ is a topology on $[0, 1]$. It is called the co-countable topology on $[0, 1]$. Now $\{ 1 \} \notin \mathcal{T}$ so that $[0, 1]$ is not closed. Hence the $\mathcal{T}$-closure of $[0, 1]$ is $[0, 1]$. However, every sequence $\{ x_n \}$ in $[0, 1]$ cannot converge to $1$ in the closure of $[0, 1]$ because $[0, 1] \setminus \{ x_1, x_2, x_3, \ldots \}$ is a $\mathcal{T}$-open neighborhood of $1$ that does not contain any term of the sequence $\{ x_n \}$. 5
To remedy the situation, we now introduce a generalization of sequence called net.

**Definitions.** (a) A directed set (or directed system) is a poset $I$ such that for every $x, y \in I$, there is $z \in I$ such that $x \preceq z$ and $y \preceq z$.

(b) A net $\{x_\alpha\}_{\alpha \in I}$ in a set $S$ is a function from a directed set $I$ to $S$ assigning every $\alpha \in I$ to a $x_\alpha \in S$.

(c) A net $\{x_\alpha\}_{\alpha \in I}$ is eventually in a set $W$ iff $\exists \beta \in I, \forall \alpha \succeq \beta$, we have $x_\alpha \in W$. A net $\{x_\alpha\}_{\alpha \in I}$ converges to $x$ (and we write $\{x_\alpha\}_{\alpha \in I} \to x$ or $x_\alpha \to x$) iff for every neighborhood $N$ of $x$, $\{x_\alpha\}_{\alpha \in I}$ is eventually in $N$.

(d) A net $\{x_\alpha\}_{\alpha \in I}$ is frequently in a set $W$ iff $\forall \beta \in I, \exists \alpha \succeq \beta$ such that $x_\alpha \in W$. We say $x$ is a cluster point of $\{x_\alpha\}_{\alpha \in I}$ iff for every neighborhood $N$ of $x$, $\{x_\alpha\}_{\alpha \in I}$ is frequently in $N$.

(e) A net $\{x_\alpha\}_{\alpha \in I}$ is a subnet of a net $\{y_\beta\}_{\beta \in J}$ iff there is a function $f : I \to J$ such that for every $\alpha \in I$, $x_\alpha = y_{f(\alpha)}$ and for every $\beta \in J$, there exists $\gamma \in I$ such that $\alpha \succeq \gamma$ implies $f(\alpha) \succeq \beta$.

**Examples.** (1) In the case $I = \mathbb{N}$ is the set of positive integers with the usual order, a net is just a sequence. In the case $I$ is an open interval $(a, b)$ of $\mathbb{R}$ with the usual order, a net in $W$ converges to $x$ is just a function from $(a, b)$ to $W$ with the left-handed limit at $b$ equals $x$. If we reverse the order on $(a, b)$, this becomes the right-handed limit at $a$ equals $x$.

(2) Convergent net need not be bounded! For example, let $I = (-\infty, 0)$ with the usual order and $x_\alpha = \alpha$. Then $x_\alpha$ converges to $0$, but $\{x_\alpha : \alpha \in I\} = (-\infty, 0)$ is unbounded!

The following theorem on topological spaces generalize the familiar theorems on uniqueness of limit, closure, continuity, cluster point and compactness for metric spaces.

**Exercises.** Prove the following statements.

(11) A topological space $X$ is Hausdorff iff every convergent net in $X$ has a unique limit.

(12) For a subset $S$ of a topological space $X$, $\overline{S} = \{x \in X : \exists \{x_\alpha\}_{\alpha \in I} in S such that x_\alpha \to x\}$.

(13) For topological spaces $X$ and $Y$, a function $f : X \to Y$ is continuous iff $f$ is continuous at every $x \in X$ iff for every $x \in X$ and $\{x_\alpha\}_{\alpha \in I}$ in $X$ with $x_\alpha \to x$, we have $f(x_\alpha) \to f(x)$. If $D$ is dense in $X$ (i.e. $\overline{D} = X$), $Y$ is Hausdorff and $f, g : X \to Y$ continuous with $|f|_D = |g|_D$, then $f = g$.

(14) $x$ is a cluster point of $\{x_\alpha\}_{\alpha \in I}$ iff $\{x_\alpha\}_{\alpha \in I}$ has a subnet converging to $x$.

(15) (Bolzano-Weierstrass Theorem) A topological space $X$ is compact iff every $\{x_\alpha\}_{\alpha \in I}$ in $X$ has a subnet converging to some $x \in X$ (equivalently, every net in $X$ has a cluster point).

For proofs, see [Be], pp. 24-26 and 35-36.

**Definition.** A topological space $X$ is sequentially compact iff every sequence in $X$ has a subsequence converging to some $x \in X$.

**Remark.** In metric spaces, compactness is the same as sequentially compactness (by the metric compactness theorem). For topological spaces, there exists a compact space that is not sequentially compact. So in such a space there is a sequence having a convergent subnet, but no convergent subsequence! Also, there is a sequentially compact set that is not compact. (See [SS], pp. 69 and 126.)

In analysis, we try to solve problems by approximations. The solutions are often some kind of limits of the approximations. So limits of convergent subsequences or convergent subnets are good candidates for the solutions. Therefore, a large part of analysis studies compactness or sequential compactness conditions.
§§2.4. **Product Topology.** We begin by asking the following

**Questions:** If we take a collection $\Omega$ of arbitrary subsets of $X$, must there exist a topology on $X$ that will contain these arbitrary subsets of $X$. We know $P(X)$ is one such topology. In fact, it is the largest such topology. Is there a smallest such topology?

To answer this question, we can first check that the intersection of any collection of topologies on $X$ is also a topology on $X$.

**Definition.** For every collection $\Omega$ of subsets of $X$, the topology $T_\Omega$ generated by $\Omega$ is the intersection of all topologies on $X$ containing $\Omega$. Hence, $T_\Omega$ is the smallest topology on $X$ containing $\Omega$.

**Exercise.** (16) Prove that $T_\Omega$ is the collection of all sets that are $\emptyset$ or $X$ or unions of sets of the form $S_1 \cap S_2 \cap \cdots \cap S_n$, where $S_1, S_2, \ldots, S_n \in \Omega$ (i.e. the set of all finite intersections of $S_i \in \Omega$ is a base of $T_\Omega$).

If we take an open interval $(a, b)$ in $\mathbb{R}$ and form $(a, b) \times \mathbb{R}$ and $\mathbb{R} \times (a, b)$, then we get “open” strips in $\mathbb{R}^2$. More generally, if $S$ is an open set in $\mathbb{R}$, then $S \times \mathbb{R}$ and $\mathbb{R} \times S$ should be “open” in $\mathbb{R}^2$. For any two topological spaces $X$ and $Y$, we would like to introduce a “product” topology on $X \times Y$ based on these ideas.

**Definitions.** For $X$ with topology $T_X$ and $Y$ with topology $T_Y$, we define the **product topology** on $X \times Y$ to be the topology $T_{X \times Y}$ generated by $\Omega = \{S \times Y : S \in T_X\} \cup \{X \times S : S \in T_Y\}$. The functions $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ defined by $\pi_X(x, y) = x$ and $\pi_Y(x, y) = y$ are called the **projection maps** onto $X$ and $Y$, respectively. Then $\Omega = \{\pi_X^{-1}(S_1) : S_1 \in T_X\} \cup \{\pi_Y^{-1}(S_2) : S_2 \in T_Y\}$ so that $\pi_X$ and $\pi_Y$ are continuous. By the exercise above, $B = \{\pi_X^{-1}(S_1) \cap \pi_Y^{-1}(S_2) : S_1 \times S_2 : S_1 \in T_X, S_2 \in T_Y\}$ is a base of $T_{X \times Y}$.

More generally, if $X_\alpha$ is a topological space with topology $T_\alpha$ for every $\alpha \in A$, then the **product topology** on their Cartesian product $X = \prod_{\alpha \in A} X_\alpha$ is the topology generated by the collection $\Omega$ of all sets of the form $\pi_\alpha^{-1}(S_\alpha)$, where $S_\alpha \in T_\alpha$ and $\pi_\alpha : X \to X_\alpha$ is the projection map $\pi_\alpha(x) = x_\alpha$ with $x_\alpha$ denoting the $\alpha$-coordinate of $x \in X$. So every $\pi_\alpha$ is continuous. A typical element in the base of the product topology is

$$\pi_\alpha^{-1}(S_\alpha) \cap \cdots \cap \pi_n^{-1}(S_n) = \bigcap_{i=1}^n \{x \in X : \pi_\alpha(x) \in S_\alpha\},$$

where $\alpha_1, \ldots, \alpha_n \in A$ and $S_\alpha \in T_\alpha_1, \ldots, S_n \in T_\alpha_n$.

In dealing with nets in product topology, we have

**Theorem.** A net $\{x_\gamma\}_{\gamma \in I}$ in $X = \prod_{\alpha \in A} X_\alpha$ converges to $x$ iff for every $\alpha \in A$, $\{\pi_\alpha(x_\gamma)\}_{\gamma \in I} \to \pi_\alpha(x)$.

**Proof.** Since the sets $\pi_\alpha^{-1}(S_\alpha) \cap \cdots \cap \pi_n^{-1}(S_n)$, where $S_\alpha \in T_\alpha$, form a base of the product topology, $\{x_\gamma\}_{\gamma \in I} \to x \iff \forall$ neighborhood $\pi_\alpha^{-1}(S_\alpha) \cap \cdots \cap \pi_n^{-1}(S_n)$ of $x$, $\exists \beta \in I$ such that $\gamma \geq \beta$ implies $\pi_\alpha(x_\gamma) \in \pi_\alpha^{-1}(S_\alpha) \cap \cdots \cap \pi_n^{-1}(S_n)$

$\iff \forall \alpha \in A, x \in \pi_\alpha^{-1}(S_\alpha) \exists \beta \in I$ such that $\gamma \geq \beta$ implies $\pi_\alpha(x_\gamma) \in \pi_\alpha^{-1}(S_\alpha)$

$\iff \forall \alpha \in A, \pi_\alpha(x) \in S_\alpha \exists \beta \in I$ such that $\gamma \geq \beta$ implies $\pi_\alpha(x_\gamma) \in S_\alpha$

$\iff \forall \alpha \in A, \{\pi_\alpha(x_\gamma)\}_{\gamma \in I} \to \pi_\alpha(x)$.

where in the second step, we take $n = 1$, $\beta_1 = \beta$ in the $\Rightarrow$ direction and take $\beta \geq \beta_i$ for $i = 1, \ldots, n$ in the $\Leftarrow$ direction.
Appendix: Proof of Zorn’s Lemma

Let us recall

**Zorn’s Lemma.** For a nonempty poset $X$, if every chain in $X$ has an upper bound in $X$, then $X$ has at least one maximal element. (The statement is also true if ‘upper’ and ‘maximal’ are replaced by ‘lower’ and ‘minimal’ respectively.)

**Proof.** (Due to H. Lenz, H. Kneser and J. Lewin independently) Assume $X$ has no maximal element. Since every chain $C$ in $X$ has an upper bound $u \in X$ and $u$ is not maximal in $X$, the set $S_u = \{ x \in X : c \in C \Rightarrow c \prec x \} \neq \emptyset$. (Here, $S_0 = X$.) By the axiom of choice, there is a function $f$ such that $f(C) \in S_C$.

We introduce two terminologies.

(a) For a chain $C$ in $X$, a set of the form $P(C, c) = \{ y \in C : y \prec c \}$ for some $c \in C$ is called an initial segment of $C$.

(b) A subset $A$ of $X$ is conforming in $X$ if (1) $A$ is well-ordered by $\leq$ and (2) for all $a \in A$, $f(P(A, a)) = a$. For example, $A = \{ f(\emptyset) \}$ is conforming because $P(A, f(\emptyset)) = \emptyset$ and so $f(P(A, f(\emptyset))) = f(\emptyset)$.

**Claim 1:** For conforming subsets $A, B$ of $X$, if $A \neq B$, then one of them is an initial segment of the other.

**Proof of claim 1.** Since $A \neq B$, either $A \subset B$ or $B \subset A$ is false, say the former, then $A \setminus B \neq \emptyset$. Let $x$ be least in $A \setminus B$, then since $a \in A$ and $a \prec x$ imply $a \in B$, we have $P(A, x) \subset B$. We will finish by showing $B = P(A, x)$. Assume $P(A, x) \neq B$. Then there is a least $y \in B \setminus P(A, x)$. Observe that for all $u \in P(B, y)$, since $u \in B$, $u \prec y$ and $y$ least in $B \setminus P(A, x)$, we get $u \in P(A, x)$. Then $u \in A$ and $u \prec x$. (*) For all $v \in A$ with $u \prec v$, since $u \prec u \prec x$, we have $v \in P(A, x) \subset B$. Next, since $\emptyset \neq A \setminus B \subseteq A \setminus P(B, y)$, so $A \setminus P(B, y)$ has a least element $z$.

We will show $P(A, z) = P(B, y)$. (First, $P(A, z) \subseteq P(B, y)$ because $w \in P(A, z)$ implies $w \in A$ and $w \prec z$, the minimality of $z$ implies $w \in P(B, y)$. For the reverse inclusion, $w \in P(B, y)$ implies $w \in B$ and $w \prec y$. The minimality of $y$ implies $w \in P(A, x)$, particularly $w \in A$. If $z \prec w$, then $z \prec y$ and setting $v = z, u = w$ in (*), we get $z \in B$. Then $z \in P(B, y)$, a contradiction. Since $w, z \in B$, so $w \leq z$. Now $w \neq z$ as $w \in P(B, y)$ and $z \notin P(B, y)$. Hence $w \prec z$, i.e. $w \in P(A, z)$. This gives us $P(B, y) \subseteq P(A, z)$.)

Next $x \in A \setminus B \subseteq A \setminus P(B, y)$ and $z$ is least in $A \setminus P(B, y)$ imply $z \leq x$. However, $z = f(P(A, z)) = f(P(B, y)) = y \in B$ and $x \notin B$. So $z \neq x$, hence $z \prec x$. Now $y = z \in P(A, x)$, contradicting the definition of $y$. Then $B = P(A, x)$. So claim 1 is proved.

**Claim 2:** Let $U = \bigcup \{ S : S$ conforming in $X \}$, $y \in U$, $A$ conforming in $X$, $x \in A$ and $y \prec x$. Then $y \in A$.

**Proof of claim 2.** Assume $y \notin A$. Now $y \in U$ imply $y \in B$ for some conforming $B$ in $X$. Then $A \neq B$. By claim 1, $A = P(B, w)$ for some $w$. Then $y \in B$, $x \in A = P(B, w)$ and $y \prec x \prec w$, so $y \in P(B, w) = A$, a contradiction. So claim 2 is proved.

**Claim 3:** $U$ is conforming.

**Proof of claim 3.** Let $x, y \in U$. There are conforming $A, B$ such that $x \in A, y \in B$. As claim 1 implies $A \subseteq B$ or $B \subseteq A$ and $A, B$ are totally ordered, so $U$ is also totally ordered.

To see $U$ is well-ordered, let $x \in G \subseteq U$, then $x$ is in some conforming $A$. If $x$ is not least in $G$, then $y \in P(G, x) \subset U$ implies $y \in A$ by claim 2. So $P(G, x) \subseteq A$ and hence $P(G, x)$ has a least element $d$. For all $g \in G$, either $g \leq x(\not\prec d)$ or $x \prec g \Rightarrow g \in P(G, x) \Rightarrow g \geq d$. So $d$ is least in $G$.

Next to get $x = f(P(U, x))$, note every $x \in U$ is in some conforming $A$. We will show $P(U, x) = P(A, x)$. First, $A \subseteq U$ implies $P(A, x) \subseteq P(U, x)$. Also $y \in P(U, x)$ implies $y \in A$ by claim 2. So $P(U, x) \subseteq P(A, x)$. Hence they are equal. Then $f(P(U, x)) = f(P(A, x)) = x$. So claim 3 is proved.

Finally, let $x = f(U) \in S_U$, then for all $u \in U$, $u \prec x$. So $x \notin U$. Note $P(U \cup \{ x \}) = U$ and for $u \in U$, $P(U \cup \{ x \}, u) = P(U, u)$. Hence $U \cup \{ x \}$ is conforming. By definition of $U$, we get $x \in U$, a contradiction. 

8
Chapter 1. Topological Vector Spaces.

In functional analysis, we deal with (usually infinite dimensional) vector spaces $X$ over $K = \mathbb{R}$ or $\mathbb{C}$ and "continuous" linear transformations between them. So we consider vector spaces with topologies and it is natural to require addition and scalar multiplication be continuous.

**Notation.** We call $K$ the **scalar field of $X$** and $K = \mathbb{R}$ or $\mathbb{C}$ for all vector spaces to be considered.

**Definitions.** A vector space $X$ with a topology is a **topological vector space** (or linear topological space) iff the topology on $X$ is a vector topology (i.e. addition $f: X \times X \to X$ defined by $f(x, y) = x + y$ and scalar multiplication $g: K \times X \to X$ defined by $g(c, x) = cx$ are continuous with respect to the topology.) For example, the indiscrete topology on $X$ is a vector topology.

**Remarks.** (1) For all $a \in X$, $T_a(x) = a + x$ is a homeomorphism. $U$ is open in $X$ iff $a + U$ is open in $X$. A linear function $h: X \to Y$ is continuous iff it is continuous at $0$ (i.e. for every neighborhood $V$ of $0$ in $Y$, $h^{-1}(V)$ is a neighborhood of $0$ in $X$). If we have a base at $0$ (or local base), which is a set $S$ of neighborhoods of $0$ such that every neighborhood of $0$ contains a member of $S$, then $B = \{a + N : a \in X, N \in S\}$ is a base. (2) For $c \neq 0$, $g_c(x) = cx$ is a homeomorphism. So $V$ is a neighborhood of $0$ implies $cV$ is a neighborhood of $0$.

**Definitions.** Let $X$ be a vector space over $K$ and $S \subseteq X$.

(1) $S$ is **convex** iff $x, y \in S$, $t \in [0, 1]$ implies $tx + (1 - t)y \in S$.

(2) $S$ is **balanced** (or **circled**) iff $x \in S$, $|c| \leq 1 \implies cx \in S$. $S$ is **absolutely convex** iff it is convex and balanced.

(3) $S$ is **absorbing** iff for every $x \in X$, there is $r > 0$ such that $0 < |c| \leq r$ implies $cx \in S$.

**Remarks.** (1) Every neighborhood $S$ of $0$ in a topological vector space is absorbing. To see this, let $x \in X$. Since the scalar multiplication $g$ is continuous and $g(0, x) = 0 \in S$, so $g^{-1}(S)$ is a neighborhood of $(0, x)$. Then there are $r > 0$ and neighborhood $U$ of $x$ such that $\{c \in K : |c| < 2r\} \times U = \pi_1^{-1}(B(0, 2r)) \cap \pi_2^{-1}(U) \subseteq g^{-1}(S)$. For $|c| \leq r$, since $x \in U$, so $cx = g(c, x) \in S$.

(2) Every neighborhood $U$ of $0$ in a topological vector space contains a balanced neighborhood of $0$. To see this, since $g(0, 0) = 0$, so there are $r > 0$ and neighborhood $V$ of $0$ such that $B(0, r) \times V \subseteq g^{-1}(U)$. So $g(\lambda, V) = \lambda V \subseteq U$ for all $|\lambda| < r$. Let $S = \bigcup_{|\lambda| < r} \lambda V$, then $S$ is a balanced neighborhood of $0$ inside $U$.

**Definitions.** Let $X, Y$ be vector spaces. For a linear function $T: X \to Y$, the **kernel** (or null space) of $T$ is $\ker T = T^{-1}(\{0\}) = \{x \in X : T(x) = 0\}$ and the **range** of $T$ is $\operatorname{ran} T = T(X) = \{T x : x \in X\}$. (Another notation for kernel of $T$ is $N(T)$ and for range of $T$ is $R(T)$.)

**Closed Kernel Theorem.** For a topological vector space $X$ and a linear function $T: X \to K$, $\ker T$ is closed if and only if $T$ is continuous. ($\mathbb{K}$ cannot be replaced by $X$ or $Y$, see [W], p. 113, ex. 3.)

**Proof.** The if direction is clear. In the only-if direction, for a $x \in X \setminus \ker T$, there is a balanced neighborhood $V$ of $0$ such that $x + V \subseteq X \setminus \ker T$, i.e. $(x + V) \cap \ker T = \emptyset$. Then $0 \notin T(x + V)$. So $T(V)$ cannot contain $-T(x) \in K$. Since $V$ is balanced, $T(V)$ is balanced in $K$. So $T(V)$ is a subset of $B(0, r) = \{z \in K : |z| < r\}$, where $r = |T(x)|$. Then for all $R > 0$, $T(B(0, R)) \subseteq B(0, r)$. So $T^{-1}(B(0, R)) \supseteq B(0, R)$. Hence, $T$ is continuous.

§1. Normed Spaces. One common type of topological vector spaces that we will deal with frequently is the family of normed linear spaces.

**Definitions.** (1) A **semi-norm** on a vector space $X$ is a function that assigns every $x \in X$ a number $||x|| \in \mathbb{R}$ satisfying (a) $||x|| \geq 0$ for all $x \in X$, (b) $||cx|| = |c||x||$ for all $c \in K, x \in X$ and (c) $||x + y|| \leq ||x|| + ||y||$ for all $x, y \in X$. It is a **norm** iff in addition to (a), (b), (c), we also have $||x|| = 0$ implies $x = 0$.

(2) A **normed space** (or normed linear space or normed vector space) is a vector space with a norm. A **Banach space** is a complete normed space (where complete means all Cauchy sequences converge). For inner
product space \( V \), define \( \|x\| = \sqrt{(x,x)} \) for all \( x \in V \). This makes \( V \) a normed space. A **Hilbert space** is a complete inner product space.

(3) For normed spaces \( X \) and \( Y \), a linear transformation from \( X \) to \( Y \) is also called a **linear operator**. In case \( Y = \mathbb{K} \), it is also called a **linear functional**. Let \( L(X,Y) \) denote the set of all continuous (equivalently, bounded) linear operators from \( X \) to \( Y \). In case \( X = Y \), we write \( L(X) \) for \( L(X,X) \). (Instead of \( L(X,Y) \), the notations \( B(X,Y) \), \( \mathcal{L}(X,Y) \) or \( B(X,Y) \) are also common.)

(4) For a topological vector space \( X \) over \( \mathbb{K} \), we write \( X^* \) for \( L(X,\mathbb{K}) \) and call it the **dual space** (or **conjugate space**) of \( X \). The elements of \( X^* \) are called the **continuous linear functionals** on \( X \).

**Examples.** (1) Let \( X \) be a normed space. For every \( x \in X \) and linear \( T : X \to \mathbb{K} \), the function \( p_T(x) = |T(x)| \) is easily checked to be a semi-norm on \( X \). It is a norm if and only if \( \ker T = \{0\} \).

(2) \( \mathbb{K}^n \) with norm \( \|(z_1, \ldots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2} \) is a Banach space. \( (\mathbb{K}^n)^* = \mathbb{K}^n \).

(3) The set of all polynomials \( P([0,1]) \) with \( \|f\| = \sup\{|f(x)| : x \in [0,1]\} \) is a normed space that is not complete. By the Weierstrass approximation theorem, \( P([0,1]) \) is dense in the set of all continuous functions \( C([0,1]) \) on \([0,1]\) with the same norm.

In general, for a compact set \( X \), let \( C(X) \) be the set of all continuous functions from \( X \) to \( \mathbb{K} \) with sup-norm \( \|f\| = \sup\{|f(x)| : x \in X\} \). Then \( C(X) \) is a Banach space. For a description of the dual of \( C(X) \), see Rudin’s *Real and Complex Analysis*, 3rd ed, p. 130.

(4) For \( 1 \leq p < \infty \) and measurable \( X \subseteq \mathbb{R} \), the Lebesgue spaces

\[
L^p(X) = \{ [f] : f \text{ measurable on } X, \|f\|_p = \left( \int_X |f|^p \, dm \right)^{1/p} < \infty \},
\]

where \([f] \) denotes the set of measurable functions equal to \( f \) almost everywhere, is a Banach space. We have \((L^p)^* = L^q\), where \( \frac{1}{p} + \frac{1}{q} = 1 \), see Rudin’s book *Real and Complex Analysis*, 3rd ed, p. 127. Such \( q \) is called the **conjugate index** to \( p \).

Also, there is \( L^\infty(X) \) consisted of all \([f] \)'s with \( f \) having finite essential sup-norm. For its dual, see Alberto Torchinsky’s book *Real Variables* p. 292.

(5) For \( 1 \leq p < \infty \), \( \ell^p = \{(a_1, a_2, a_3, \ldots) : a_i \in \mathbb{K}, \|(a_1, a_2, a_3, \ldots)\|_p = \left( \sum |a_i|^p \right)^{1/p} < \infty \} \) is a Banach space. The dual of \( \ell^p \) is \( \ell^q \), where \( q \) is the conjugate index of \( p \). (Instead of \( \ell^p \), the notation \( \ell_p \) is also common.)

(6) \( \ell^\infty = \{(a_1, a_2, a_3, \ldots) : a_i \in \mathbb{K}, \sup\{|a_i| : i = 1, 2, 3, \ldots\} < \infty \} \) is a Banach space. Its dual is the dual of \( L^\infty(\mathbb{N}) \).

The spaces

\[
c = \{(a_1, a_2, a_3, \ldots) : a_i \in \mathbb{K}, \lim_{i \to \infty} a_i \in \mathbb{K} \} \quad \text{and} \quad c_0 = \{(a_1, a_2, a_3, \ldots) : a_i \in \mathbb{K}, \lim_{i \to \infty} a_i = 0 \}
\]

are closed vector subspaces of \( \ell^\infty \). Hence, they are Banach spaces with the same norm as \( \ell^\infty \).

(7) Let \( X, Y \) be normed spaces. For \( T \in L(X,Y) \), define \( \|T\| = \sup\{\|T(x)\| : \|x\| \leq 1\} \). It is easy to check that \( L(X,Y) \) is a normed space.

(8) Let \( X, Y \) be normed spaces. For \( 1 \leq p < \infty \), we may define \( X \oplus Y = \{(x,y) : x \in X, y \in Y\} \) with \( \|(x,y)\|_p = \left( \|x\|^p + \|y\|^p \right)^{1/p} \). It is easy to check that \( X \oplus Y \) is a normed space with \( \| \cdot \|_p \) as norm. It is also possible to use \( \|(x,y)\|_\infty = \max\{|x|, |y|\} \) as norm. All these norms are equivalent. We called \( X \oplus Y \) the **direct sum** of \( X \) and \( Y \). If \( X, Y \) are Banach spaces, then \( X \oplus Y \) is also a Banach space. For Hilbert spaces \( X \) and \( Y \), the direct sum \( X \oplus Y \) with the inner product given by \( \langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle \) inducing the norm \( \|(x,y)\|_2 = (\|x\|^2 + \|y\|^2)^{1/2} \) is a Hilbert space.

The **projection map** \( P_X : X \oplus Y \to X \) defined by \( P_X(x,y) = x \) is continuous since \( \|x\| \leq \|(x,y)\|_p \) and similarly, the projection map \( P_Y : X \oplus Y \to Y \) defined by \( P_Y(x,y) = y \) is continuous.
(9) Let $N$ be a closed vector subspace of a normed space $X$. For $x \in X$, we define $[x] = x + N = \{x+n : n \in N\}$ and $X/N = \{[x] : x \in X\}$. For $c \in \mathbb{R}$ and $x, y \in Y$, defining $[x] + [y] = [x + y]$ and $c[x] = [cx]$ shows $X/N$ is a vector space with $[0] = 0 + N = N$.

Next define $|[x]| = \inf\{|x - n| : n \in N\}$. We have $|[x]| = 0$ implies there is a sequence $\{n_k\}$ in $N$ such that $|x - n_k| \to 0$ so that $n_k \to x \in \overline{N} = N$ and $[x] = [0]$. It is easy to see that this makes $X/N$ a normed space. We call $X/N$ the quotient normed space of $X$ by $N$, and $\|\cdot\|$ the quotient norm. The linear surjection $\pi_N : X \to X/N$ defined by $\pi_N(x) = [x]$ is called the quotient map. It is continuous since $|[x]| = \inf\{|x - n| : n \in N\} \leq \|x\|$. Also, $\pi_N(B(0, 1)) = B(0, 1)$ implies it maps open sets to open sets.

**Theorem.** If $N$ is a closed vector subspace of a Banach space $X$, then $X/N$ is also a Banach space.

**Proof.** Recall that a normed space is complete iff every absolutely convergent series converges in the space. Suppose $\sum_{k=1}^{\infty} \|x_k\| < \infty$. By infimum property, for every $k$, there exists $n_k \in N$ such that $\|x_k - n_k\| \leq 2\inf\{|x_k - n| : n \in N\} = 2\|[x_k]\|$. Then $\sum_{k=1}^{\infty} \|x_k - n_k\| < \infty$. Since $X$ is complete, this implies $\sum_{k=1}^{\infty} (x_k - n_k)$ converges to some $x \in X$. Using $\|w\| \leq \|w\|$ for all $w \in X$, we have

$$\left\| \sum_{k=1}^{m} [x_k] - [x] \right\| = \left\| \sum_{k=1}^{m} x_k - x \right\| = \left\| \sum_{k=1}^{m} x_k - x - \sum_{k=1}^{m} n_k \right\| \leq \sum_{k=1}^{m} \|x_k - n_k - x\| \to 0 \text{ as } m \to \infty.$$  

**Remarks.** The same reasoning also show that if $E$ is a subspace of a Banach space $X$ such that $E + N$ is closed (hence complete) in $X$, then $(E + N)/N$ is complete, hence closed in $X/N$.

**Definition.** For a closed vector subspace $N$ of a Banach space $X$, define the codimension of $N$ to be $\text{codim } N = \text{dim } X/N$.

**Remark.** In [RS], pp. 102-103, there is a nice functional analysis proof of the Tietze extension theorem on compact spaces using quotient spaces.

§2. Locally Convex Spaces. Occasionally, we will come across vector spaces $X$ that have many important (semi-)norms like those of the form $\|T(x)\|$, where $T : X \to \mathbb{K}$ is linear. Then we may want vector topologies on the vector spaces so that all these (semi-)norms are continuous.

**Definition.** $X$ is a locally convex space iff $X$ is a topological vector space such that every neighborhood of 0 contains a convex neighborhood of 0.

**Theorem.** (a) Let $p(x) = \|x\|$ be a seminorm on a vector space $X$. Then the unit balls $V_1 = \{x : p(x) < 1\}$ and $V_2 = \{x : p(x) \leq 1\}$ are absorbing and absolutely convex.

(b) A seminorm $p(x)$ on a topological vector space $X$ is continuous iff $V_1 = \{x : p(x) < 1\}$ is a neighborhood of 0 in $X$ iff $V_2 = \{x : p(x) \leq 1\}$ is a neighborhood of 0 in $X$.

(c) Let $\mathcal{P}$ be a family of seminorms on a vector space $X$. For every $p \in \mathcal{P}$, let $V(p) = \{x : p(x) < 1\}$. The collection

$$U = \{r_1 V(p_1) \cap \cdots \cap r_n V(p_n) : r_1, \ldots, r_n > 0, \; p_1, \ldots, p_n \in \mathcal{P} \}$$

is a base at 0 of a topology that makes $X$ into a locally convex space. Furthermore, it is the weakest vector topology on $X$ for which all seminorms in $\mathcal{P}$ are continuous.

**Proof.** See [TL], pp. 105-107.
**Remarks.** (1) The topology given in (c) is called the **topology generated by the family** $P$ **of seminorms**.
(2) In the case $P$ is consisted of exactly one norm, then we get a normed space. So all theorems on locally convex spaces apply to normed spaces!

**Theorem.** (a) Let $X$ be a locally convex space whose topology is generated by a family $P$ of seminorms. $X$ is Hausdorff iff $P$ is separating (i.e. for each nonzero $x \in X$, there is $p \in P$ such that $p(x) \neq 0$).
(b) A topological vector space $X$ is a locally convex space iff there exists a family of seminorms that generates the topology on $X$.

**Proof.** See [TL], p. 107 for (a) and p. 113 for (b).

**Definition.** A set $S$ in a topological vector space $X$ is **bounded** iff for every neighborhood $N$ of $0$, there is $r > 0$ such that $S \subseteq rN$.

**Theorem.** Let $X$ be a locally convex space whose topology is generated by a family $P$ of seminorms.
(a) A set $W$ is bounded in $X$ iff for every $p \in P$, $p(W)$ is bounded in $K$.
(b) A net $\{x_\alpha\}_{\alpha \in I} \to x$ in $X$ iff for every $p \in P$, $\{p(x_\alpha - x)\}_{\alpha \in I} \to 0$. (Then $|p(x_\alpha) - p(x)| \leq p(x_\alpha - x) \to 0$.)

**Proof.** (a) $W$ is bounded $\iff \forall p_1, \ldots, p_n \in P, r_1, \ldots, r_n > 0, \exists r > 0$ such that $W \subseteq r \bigcap_{i=1}^{n} \{x : p_i(x) < r_i\}$

$$\iff \forall p_i \in P, \exists R_i > 0$$ such that $\forall x \in W, p_i(x) < R_i$

$$\iff \forall p \in P, \{p(W)\}$$ is bounded in $K$,

where in the second step, we take $n = 1, R_1 = rR_1$ in the $\Rightarrow$ direction and take $r > R_i/r_i$ for $i = 1, \ldots, n$ in the $\Leftarrow$ direction.

(b) $\{x_\alpha\}_{\alpha \in I} \to x$ $\iff \{x_\alpha - x\}_{\alpha \in I} \to 0$

$$\iff \forall p_1, \ldots, p_n \in P, r_1, \ldots, r_n > 0, \exists \beta \in I$$ such that

$$\alpha \geq \beta$$ implies $x_\alpha - x \in \bigcap_{i=1}^{n} \{y : p_i(y) < r_i\}$

$$\iff \forall p_i \in P, r_i > 0, \exists \beta_i \in I$$ such that $\alpha \geq \beta_i$ implies $x_\alpha - x \in \{y : p_i(y) < r_i\}$

$$\iff \forall p \in P, \{p(x_\alpha - x)\}_{\alpha \in I} \to 0.$$  

**Questions:** Why are we interested in locally convex spaces? Why are normed spaces not good enough?

(1) Some important classes of functions in analysis, such as the collection of distributions or generalized functions is not a normed space. They can be topologized by seminorms.

(2) In analysis, we solve many problems by taking limit. Very often we consider bounded sequences and try to extract convergent subsequences or subnets to get a limit point. For an infinite dimensional normed space, an application of the Riesz lemma showed the closed unit ball is not compact. So bounded sequences on normed spaces may not have convergent subsequences or subnets!

For a normed space $X$, there is a weakest vector topology $w$ on $X$ that makes all elements of $X^*$ continuous. We simply take $P = \{|f| : f \in X^*\}$ and apply the theorems above. This topology on $X$ is called the **weak topology** on $X$. Then $X$ with the weak topology is a locally convex space.

Similarly, on a dual space $X^* = L(X, K)$ (which is a normed space), for each $x \in X$, we can define $i_x : X^* \to K$ by $i_x(y) = y(x)$. The inequality $|i_x(y') - i_x(y)| = |y'(x) - y(x)| \leq \|y' - y\|\|x\|$ implies $i_x \in (X^*)^* = X^{**}$. We can take $P = \{|i_x| : x \in X\}$ to generate a topology $w^*$ on $X^*$ so that all $i_x$ are continuous. This topology $w^*$ on $X^*$ is called the **weak-star topology** on $X^*$.

The important fact is that Banach and Alaoglu proved that the closed unit ball of $X^*$ is $w^*$-compact, i.e. compact in the weak-star topology. So bounded sequences on dual spaces have $w^*$-cluster points for solving analysis problems. We will prove this later.
Chapter 2. Basic Principles.

§1. Consequences of Baire’s Category Theorem. In this and next sections, we will study important principles about linear operators between topological vector spaces. The four pillars of functional analysis are the open mapping theorem, the closed graph theorem, the uniform boundedness principle and the Hahn-Banach theorem. They have many applications in different branches of mathematics. We will cover the first three of these in this section and the last one in the next section.

Definition. For topological spaces $X$ and $Y$, $T : X \to Y$ is open iff $U$ open in $X$ implies $T(U)$ open in $Y$.

Remarks. (1) In checking $T : X \to Y$ is open, it is enough to check $T(U)$ is open for $U$’s in a base of $T_X$. Then $T$ open follows from $T(\cup a U_a) = \cup a T(U_a)$. For example, every quotient map $\pi : X \to X/N$ of normed spaces is open since $\pi(B(a, r)) = B([a], r)$. Also, a projection $\pi_\beta : \prod a \rightarrow X_\beta$ is an open map since for open sets $S_\alpha$ in $X_\alpha$, $\pi_\beta(\prod a^{-1}(S_\alpha)) = S_\alpha$ or $X_\beta$ depending if $\beta = \alpha_i$ for some $i$ or not.

(2) An open map may not take closed sets to closed sets. To see this, let $X = P([0, 1])$ and $Y = C([0, 1])$ be the sets of all polynomials and continuous function on $[0, 1]$ with sup-norm, respectively. Then $V = \{(f, f) : f \in P([0, 1])\}$ is closed in $X \times Y$ because $(f_n, f_n) \rightarrow (f, g)$ in $X \times Y$ implies $f_n \rightarrow f$ in $X$ and $f_n \rightarrow g$ in $Y$, hence, by uniqueness of limit in $Y$, $f = g$ and so $(f, g) \in V$. The projection map $\pi_Y : X \times Y \to Y$ is open, but $\pi_Y(V) = X$ is not closed in $X$ since $X$ is a proper subset of $Y = X$.

If a vector subspace $M$ contains some $B(a, r)$ in a normed space $Y$, then $M = \text{span}\{B(a, r) - a\} = Y$. So if linear $T : X \to Y$ is open (or just $M = T(X)$ contains a ball of $Y$), then $T$ is surjective. Is there any converse?

Lemma. Let $X$ and $Y$ be normed spaces. A linear function $T : X \to Y$ is open if and only if there exist $r, r’ > 0$ such that $T(B(0, r)) \supseteq B(0, r’)$. 

Proof. If $T$ is open, then $T(B(0, r)) \supseteq B(0, r’)$ and contains $0$. So $T(B(0, r)) \supseteq B(0, r’)$ for some $r’ > 0$.

Next, if $T(B(0, r)) \supseteq B(0, r’)$, then since every open $U$ in $X$ is a union of $B(a, r_a) = a + r_a B(0, r)$, $T(U) = T\left( \bigcup a U \right) = \bigcup a T(a, r_a) = T(B(0, r))$ is open. So $T$ is open.

Lemma. Let $X$ be a Banach space, $Y$ a normed space and $T \in L(X, Y)$. If $T(B(0, r)) \supseteq B(0, r’)$, then $T(B(0, r)) \supseteq B(0, r’)$.

Proof. Let $y \in B(0, r’)$ and choose $c$ such that $\|y\|/r’ < c < 1$. Let $\varepsilon \in (0, 1 - c)$. Since $y \in B(0, r’)$, then $x_1 \in B(0, r)$ such that $\|y - T x_1\| < \varepsilon r’$. So $y - T x_1 \in \varepsilon c B(0, r’) \subseteq T(\varepsilon c B(0, r’))$. Iterating this, we get by induction a sequence $\{x_n\} \in X$ such that $x_n \in \varepsilon^{n-1} c B(0, r)$ and $y - T x_1 - \cdots - T x_n \in \varepsilon^n c B(0, r’)$.

Now $\sum_{n=1}^\infty \|x_n\| \leq \frac{cr}{1 - \varepsilon} < r$. Since $X$ is complete, $\sum_{n=1}^\infty x_n = x$ for some $x \in B(0, r)$. Since $T$ is continuous, $\|y - T x\| = \lim_{n \to \infty} \|y - T x_1 - \cdots - T x_n\| \leq \lim_{n \to \infty} \varepsilon^n r’ = 0$. Then $y = T x \in T(B(0, r))$.

Open Mapping Theorem. For Banach spaces $X, Y$ and $T \in L(X, Y)$, if $T$ is surjective, then $T$ is open.

Proof. Let $U_n = B(0, n)$ in $X$. Since $T(X) = T(\bigcup_{n=1}^\infty U_n) = \bigcup_{n=1}^\infty T(U_n)$ is of the second category in $Y$ by the Baire category theorem, there is $n$ such that $T(U_n)$ contains an open ball, say $B(T a, r) = T a + B(0, r)$, where $a \in U_n$. Then $B(0, r) = -T a + B(T a, r) \subseteq -T a + T(U_n) \subseteq T(U_{2n})$. By the lemmas above, $B(0, r) \subseteq T(U_{2n})$ and $T$ is open.

Remark. Let $X$ be a Banach space, $Y$ be a normed space and $T \in L(X, y)$. The proof above actually showed if $T(X)$ is of second category in $Y$, then $T$ is open and surjective.
**Definitions.** Let $X, Y$ be normed spaces. $T \in L(X, Y)$ is **invertible** if $T$ is bijective and $T^{-1} \in L(Y, X)$. $X$ and $Y$ are **isomorphic** if there is an invertible $T \in L(X, Y)$. (Such an invertible $T$ is called an **isomorphism** between $X$ and $Y$. In that case, there exist $c_1, c_2 > 0$ such that for all $x \in X$, $c_1\|x\| \leq \|Tx\| \leq c_2\|x\|$.)

**Inverse Mapping Theorem.** For Banach spaces $X$ and $Y$, if $T \in L(X, Y)$ is bijective, then $T^{-1} \in L(Y, X)$.

**Proof.** For $T \in L(X, Y)$, $T$ bijective is equivalent to $T$ injective and open (by the open mapping theorem). For all open $U$ in $X$, $(T^{-1})^{-1}(U) = T(U)$ is open in $Y$. So $T^{-1}$ is continuous.

**Isomorphism Theorem.** For normed spaces $X$ and $T \in L(X, Y)$, the linear function $\hat{T} : X/\ker T \to Y$ defined by $\hat{T}([x]) = T(x)$ is bounded and $\|\hat{T}\| = \|T\|$. In case $X$ and $Y$ are Banach spaces, if $T \in L(X, Y)$ is surjective, then $\hat{T}$ is an isomorphism and $X/\ker T$ is isomorphic to $Y$ as Banach spaces.

**Proof.** For all $n \in \ker T$, $\|\hat{T}([x])\| = \|T(x)\| = \|T(x-n)\| \leq \|T\||x-n\|$. Taking infimum over all $n \in \ker T$, we get $\|\hat{T}([x])\| \leq \|T\||[x]\|$. So $\hat{T}$ is bounded and $\|\hat{T}\| \leq \|T\|$. Next, $\|T(x)\| = \|\hat{T}([x])\| \leq \|\hat{T}\||[x]\| \leq \|T\||x||$. implies $\|T\| \leq \|\hat{T}\|$. Therefore, $\|\hat{T}\| = \|T\|$.

In case $X$ and $Y$ are Banach spaces, if $T$ is surjective, then $\hat{T}$ is bijective. By the inverse mapping theorem, $\hat{T}$ is an isomorphism.

**Remarks.** Using the inverse mapping theorem, it can be showed that there exists a complex sequence with limit zero such that it is not the Fourier coefficient sequence of a $L^1$ function on the unit circle. See applications at the end of the chapter.

**Definition.** Let $X, Y$ be normed spaces. $T \in L(X, Y)$ is **bounded below** if there exists $c' > 0$ such that for all $x \in X$, $\|Tx\| \geq c'\|x\|$.

**Remarks.** (1) Taking $u = x/\|x\|$, the inequality is the same as $\inf\{\|T(u)\| : \|u\| = 1\} > 0$. So $T$ is not bounded below if there is $\|u_n\| = 1$ and $T(u_n) \to 0$.

(2) If $T \in L(X, Y)$ is bounded below and $W$ is a complete subset of $X$, then $T(W)$ is also a complete subset in $Y$ (since for $x_n \in W$, $\{Tx_n\}$ Cauchy implies $(x_n)$ Cauchy, hence by completeness of $W$, $x_n \to x$ for some $x \in W$ and by continuity of $T$, $Tx_n \to Tx = T(x) \in T(W)$). In case $X$ and $Y$ are Banach spaces, $T$ bounded below and $W$ closed subset in $X$ imply $T(X)$ closed in $Y$.

**Lower Bound Theorem.** Let $X$ be a Banach space and $Y$ be a normed space. For $T \in L(X, Y)$, the following are equivalent:

(a) $T$ is bounded below,
(b) $T$ is injective and $T(X)$ is complete (hence closed in $Y$),
(c) $T$ has a continuous inverse $T^{-1} : L(X, Y)$.

**Proof.** (a) $\Rightarrow$ (b) If $T$ is bounded below, then $T(x) = 0$ implies $x = 0$, so $T$ is injective. By remark (2), $T(X)$ is complete (hence closed in $Y$).

(b) $\Rightarrow$ (c) This follows immediately from the inverse mapping theorem.

(c) $\Rightarrow$ (a) If $T^{-1} \in L(T(X), X)$, then $\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\|\|Tx\|$ for all $x \in X$ and we can take $0 < c' < 1/\|T^{-1}\|$.

**Remarks.** Let $X$ and $Y$ be Banach spaces. $T \in L(X, Y)$ is invertible if and only if $T$ is bounded below and has a dense range. For injective $T \in L(X, Y)$, $T$ has a closed range iff $T$ is bounded below.

For the next theorem, we introduce the

**Definition.** For topological spaces $X$ and $Y$, $T : X \to Y$ is **closed** if its graph $\Gamma(T) = \{(x, Tx) : x \in X\}$ is closed in $X \times Y$ (i.e. if $(x_\alpha, Tx_\alpha) \to (x, y) \in \Gamma(T)$, then $y = Tx$ so that $(x, y) \in \Gamma(T)$).
Recall that the projection maps \( \pi_1 : X \times Y \rightarrow X \) defined by \( \pi_1(x,y) = x \) and \( \pi_2 : X \times Y \rightarrow Y \) defined by \( \pi_2(x,y) = y \) are open and continuous.

**Closed Graph Theorem.** Let \( X, Y \) be Banach spaces and \( T : X \rightarrow Y \) be linear. If \( T \) is closed, then \( T \) is continuous.

**Proof.** Since \( X \) and \( Y \) are complete, so \( X \times Y \) is complete. Since \( \Gamma(T) = \{ (x, Tx) : x \in X \} \) is closed in \( X \times Y \), \( \Gamma(T) \) is complete. Note \( \pi_1|_{\Gamma(T)} : \Gamma(T) \rightarrow X \) is bijective. Also, \( \pi_1 \) continuous implies \( \pi_1|_{\Gamma(T)} \in L(\Gamma(T), X) \).

By the inverse mapping theorem, \( \pi_1|_{\Gamma(T)}^{-1} \in L(X, \Gamma(T)) \). Therefore, \( T = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1} \in L(X, Y) \).

**Remarks.** (1) In using the closed graph theorem, note \( \Gamma(T) \) is closed in \( X \times Y \) iff for every sequence \( (x_n, Tx_n) \in \Gamma(T) \) converging to \((x, y) \in X \times Y \), we have \( y = Tx \) (so that \( (x, y) \in \Gamma(T) \)).

(2) For every Hausdorff space \( Y \), every continuous \( T : X \rightarrow Y \) is closed because \( x_n \rightarrow x \), \( Tx_n \rightarrow y \) implies \( y = Tx \) by continuity and uniqueness of limit.

**Exercises.** (1) Let \( X \) be a vector space equipped with two complete norms \( \| \cdot \|_1 \) and \( \| \cdot \|_2 \). If there exists \( c > 0 \) such that for all \( x \in X \), \( \|x\|_1 \leq c\|x\|_2 \), prove that there exists \( c' > 0 \) such that for all \( x \in X \), \( \|x\|_2 \leq c'|x|_1 \). This means the norms are equivalent.

(2) (Hellinger-Toeplitz Theorem) Let \( H \) be a Hilbert space and \( T : H \rightarrow H \) be a linear transformation such that for all \( x, y \in H \), \( (x, Ty) = (Tx, y) \). Prove that \( T \) is bounded. (This theorem has important consequence in mathematical physics. See [RS], p. 84)

**Application.** See [Fr], pp. 145-149 or [Y], pp. 80-81 for applications of the closed graph theorem to PDE.

**Uniform Boundedness Principle (or Resonance Theorem).** Let \( X, Y \) be normed spaces, \( A \subseteq L(X, Y) \) and \( S \) be of the second category in \( X \). If \( \{ \|Tx\| : T \in A \} \) is bounded for all \( x \in S \), then \( \{ \|T\| : T \in A \} \) is bounded. (For Banach space \( X \), if \( \{ \|Tx\| : T \in A \} \) is bounded for all \( x \in X \), then \( \{ \|T\| : T \in A \} \) is bounded.)

**Proof.** Note \( S_n = \{ x \in X : \sup\{\|Tx\| : T \in A \} \leq n \} = \bigcap_{T \in A} \{ x \in X : \|Tx\| \leq n \} \) is closed. Since \( S = \bigcup_{n=1}^{\infty} S_n \), by the Baire category theorem, there is a \( S_n \) containing some ball \( B(x, r) \). Hence \( S_n \supseteq B(x, r/2) = x + B(0, r/2) \).

For every \( y \|y\| \leq 1 \), since \( x \in S_n \) and \( x + ry \in B(x, r) \subseteq S_n \), so for all \( T \in A \),

\[
\|Ty\| = \frac{\|T(ry)\|}{r} \leq \frac{\|T(ry)\| + \|Tx\|}{r} \leq \frac{2n}{r}.
\]

Therefore, for every \( T \in A, \|T\| \leq 2n/r \).

**Theorem (Banach-Steinhaus).** Let \( X \) be a Banach space, \( Y \) be a normed space and \( T_n \in L(X, Y) \).

(a) If all \( x \in X \), \( \{T_nx\} \) converges in \( Y \), then \( Tx = \lim_{n \rightarrow \infty} T_nx \in L(X, Y) \) with \( \|T\| \leq \liminf \|T_n\| \).

(b) Suppose there is \( C > 0 \) such that \( \|T_n\| \leq C \) for \( n = 1, 2, 3, \ldots \). For \( T_0 \in L(X, Y) \), the vector subspace \( M = \{ x \in X : \lim_{n \rightarrow \infty} T_nx = T_0x \} \) is a closed in \( X \). If \( M \) is dense or of the second category in \( X \), then \( M = X \) (i.e. \( T_n \) converges pointwise on \( X \) to \( T_0 \)).

**Proof.** (a) For all \( x \in X \), \( \{T_n(x)\} \) converges implies it is bounded. By the uniform boundedness principle, \( \sup\{\|T_n\| : n = 1, 2, 3, \ldots \} < \infty \). Now there is a subsequence \( \{\|T_{n_i}\|\} \) converging to \( c = \lim_{n \rightarrow \infty} \|T_n\| \). Then \( \|Tx\| = \lim_{i \rightarrow \infty} \|T_{n_i}x\| \leq \lim_{i \rightarrow \infty} \|T_{n_i}\| \|x\| = c\|x\| \), which implies \( \|T\| \leq c \).

(b) For every \( x \in M \) and \( \varepsilon > 0 \), there is \( y \in M \) such that \( \|x - y\| < \varepsilon/(2C + 2\|T_0\|) \). Since \( y \in M \), so \( T_ny \) converges to \( T_0y \). Hence, there is \( N \) such that \( n \geq N \) implies \( \|T_ny - T_0y\| < \varepsilon/2 \). Then

\[
\|T_nx - T_0x\| \leq \|T_nx - T_ny\| + \|T_ny - T_0y\| + \|T_0y - T_0x\| \leq (\|T_n\| + \|T_0\|)\|x - y\| + \varepsilon/2 < \varepsilon.
\]

15
Using the uniform boundedness principle, it can be proved that there exists a 2 M

(1) Seminorms are Minkowski functionals.

We will discuss these one at a time.

Hahn-Banach theorem

Definitions.

We define a function F : A → B as an extension of another function f : C → B if A ⊆ B and f(x) = F(x) for all x ∈ C, equivalently graph of F contains graph of f (in short F = f|C). We say F is a linear extension of f when A, B, C are vector spaces and F, f are linear.

Examples. (1) Seminorms are Minkowski functionals.

(2) In [TL], pp. 112-113, Theorem 12.1 asserts that for every convex absorbing set U, pu(x) = inf{t > 0 : x ∈ tU} is a Minkowski functional. This is called the Minkowski functional of U. It is a seminorm if U is balanced. Theorem 12.2 asserts that \(\{x : pu(x) < 1\} \subseteq U \subseteq \{x : pu(x) \leq 1\}\). Theorem 12.3 asserts that if U is an open, absolutely convex neighborhood of 0, then U = \{x : pu(x) < 1\}.

Real Hahn-Banach Theorem. Let Y be a vector subspace of a vector space X over \(\mathbb{R}\), p be a Minkowski functional on X and f : Y → \(\mathbb{R}\) be a linear function such that f(x) ≤ p(x) for all x ∈ Y. Then f has a linear extension F : X → \(\mathbb{R}\) such that F(x) ≤ p(x) for all x ∈ X.

Proof. Consider the collection S of all (Z, f_Z), where Z is a vector subspace of X containing Y and there exists a linear extension f_Z of f and f_Z(x) ≤ p(x) for all x ∈ Z. Since (Y, f) ∈ S, S ≠ ∅. Partial order the elements of S by inclusion (i.e. \((Z_0, f_{Z_0}) \subseteq (Z_1, f_{Z_1})\) if \(Z_0 \subseteq Z_1\) and \(f_{Z_1}|_{Z_0} = f_{Z_0}\)). If C is a chain in S, then we see that C has an upper bound in S. Hence, by Zorn’s lemma, S has a maximal element \((M, f_M)\).

Assume M ≠ X. Let x ∈ X \ M. Consider \(Z = \operatorname{span}(M \cup \{x\}) = M + \mathbb{R}x\). Now f_M be a linear extension of f and f_M(x) ≤ p(x) for all x ∈ M. For every a, b ∈ M,

\[f_M(a) + f_M(b) = f_M(a + b) \leq p(a + b) \leq p(a - x) + p(x + b).\]
The complexification lemma is useful in reducing problems to the case of vector spaces over $\mathbb{C}$. Specifically, it states that if $f$ is a linear functional on a vector space $X$ over $\mathbb{R}$, then we can extend $f$ to a functional $f_Z$ on the complexification $X \otimes \mathbb{C}$ in a way that preserves linearity and boundedness.

**Proof.** Let $X$ be a vector space over $\mathbb{R}$, and let $f : X \to \mathbb{R}$ be a linear functional. Then $f_Z : X \otimes \mathbb{C} \to \mathbb{C}$, defined by $f_Z(x \otimes 1) = f(x)$, is a linear functional on $X \otimes \mathbb{C}$.

The extension $f_Z$ is uniquely determined by the requirement that $f_Z(x \otimes 1) = f(x)$ for all $x \in X$. This follows from the uniqueness of linear extensions of vector spaces.

**Complexification Lemma.** Let $X$ be a vector space over $\mathbb{C}$. If $U : X \to \mathbb{R}$ is linear (considering $X$ as a vector space over $\mathbb{R}$), then $F : X \to \mathbb{C}$ defined by $F(x) = U(x) - iU(ix)$ is linear (considering $X$ as a vector space over $\mathbb{C}$).

**Proof.** For $c \in \mathbb{R}$, $x, y \in X$,

$$U(x + y) = U(x) + U(y), U(cx) = cU(x) \quad \text{implies} \quad F(x + y) = F(x) + F(y), F(cx) = cF(x).$$

Also, $F(ix) = U(ix) - iU(-ix) = i(U(x) - iU(ix)) = iF(x)$. Therefore, $F$ is linear (considering $X$ as a vector space over $\mathbb{C}$).

**Complex Hahn-Banach Theorem.** Let $Y$ be a vector subspace of a vector space $X$ over $\mathbb{C}$, $p$ be a seminorm on $X$ and $f : Y \to \mathbb{C}$ be a linear function such that $|f(x)| \leq p(x)$ for all $x \in Y$. Then $f$ has a linear extension $F : X \to \mathbb{C}$ and $|F(x)| \leq p(x)$ for all $x \in X$.

**Proof.** Let $u = \text{Re} f$ and $v = \text{Im} f$. Since $f(ix) = if(x)$, we have $u(ix) + iv(ix) = iu(x) - v(x)$ so that $\text{Im} f(x) = v(x) = -u(ix)$. Since for all $x \in Y$, $u(x) \leq |f(x)| \leq p(x)$, by the last theorem, there exists a linear extension $U : X \to \mathbb{R}$ of $u$ (considering $X$ as a vector space over $\mathbb{R}$) and $U(x) \leq p(x)$ for all $x \in X$.

By the complexification lemma, $F : X \to \mathbb{C}$ defined by $F(x) = U(x) - iU(ix)$ is linear (considering $X$ as a vector space over $\mathbb{C}$). $F$ extends $f$ because for every $x \in Y$,

$$F(x) = U(x) - iU(ix) = u(x) - iu(ix) = \text{Re} f(x) + i\text{Im} f(x) = f(x).$$

If $F(x) = 0$, then $|F(x)| = 0 = U(x) \leq p(x)$. For $F(x) \neq 0$, let $c = |F(x)|/F(x)$, then since $p$ is a seminorm,

$$|F(x)| = cF(x) = F(cx) = \text{Re} F(cx) = U(cx) \leq p(cx) = |c|p(x) = p(x).$$

**Remark.** The complexification lemma is useful in reducing problems to the case of vector spaces over $\mathbb{R}$.

**Theorem (Hahn-Banach).** Let $X$ be a normed space and $Y$ be a vector subspace of $X$.

(a) For every $f \in Y^*$, there exists an extension $F \in X^*$ of $f$ such that $\|F\| = \|f\|$.

(b) Let $x \in X$. We have $x \notin \overline{Y}$ if and only if there exists $F \in X^*$ such that $\|F\| = 1$, $F \equiv 0$ on $Y$ and $F(x) = d(x, Y) = \inf \{\|x - y\| : y \in Y\} \neq 0$. In particular, $\overline{Y} = X$ if and only if $F \in X^*$ with $F \equiv 0$ on $Y$ implies $F \equiv 0$ on $X$.

(c) If $X \neq \{0\}$, then for every $x \in X$, there exists $F \in X^*$ with $\|F\| = 1$ and $F(x) = \|x\|$. Such a function $F$ is called a support functional at $x$. 

17
Now it suffices to prove the case $\beta < \alpha$. Define $\delta = |\alpha - \beta|$ for all $\alpha, \beta \in \mathbb{K}$, then there is $\epsilon > 0$ such that $|\alpha - \beta| < \epsilon$. Thus, for all $\epsilon > 0$, there is a neighborhood of 0. Let $x_0 = \epsilon x$, $g^{-1} \left( U \right) \cap N_{x_0} \left( (0, x_0) \right)$ contains a neighborhood $(-\epsilon, \epsilon) \times N_{x_0}$ of $(0, x_0)$. This implies $t x_0 \in U$ for $t \in (-\epsilon, \epsilon)$. Now $(f(a) - \epsilon, f(a) + \epsilon) = (f(a) + (-\epsilon, \epsilon)) = \{ f(a + t x_0) \colon t \in (-\epsilon, \epsilon) \} \subseteq f(a + U) = f(A)$. So $f(A)$ is open.

**Separation Theorem.** Let $A, B$ be disjoint, nonempty convex subsets of a topological vector space $X$.

(a) If $A$ is open, then there is $f \in X^*$ such that for all $a \in A$, $\text{Re} f(x) < \inf \{ \text{Re} f(y) : y \in B \}$.

(b) (V. L. Klee, 1951) If $A$ is compact, $B$ is closed and $X$ is locally convex, then there is $f \in X^*$ such that $\sup \{ \text{Re} f(x) : x \in A \} < \inf \{ \text{Re} f(y) : y \in B \}$.

**Proof.** It suffices to prove the case $\mathbb{K} = \mathbb{R}$. (Then for the case $\mathbb{K} = \mathbb{C}$, we may regard $X$ as a vector space over $\mathbb{R}$ and keep the same topology so that it is a topological vector space over $\mathbb{R}$. Then apply the case $\mathbb{K} = \mathbb{R}$ and use the complexification lemma to get the desired complex linear functional. This complex linear functional is continuous because its real and imaginary parts are continuous.)

(a) Fix $a_0 \in A$ and $b_0 \in B$. Let $x_0 = b_0 - a_0$, then $C = A - B + x_0 = \bigcup_{b \in B} (A - b + x_0)$ is an open convex neighborhood of 0. Let $p(x) = \inf \{ t > 0 : x \in t C \}$ be the Minkowski functional of $C$, then $\{ x : p(x) < 1 \} \subseteq C$ (because $p(x) < 1$ implies there is $t \in [p(x), 1)$ such that $x \in t C \subseteq C$ and $C \subseteq \{ x : p(x) \leq 1 \}$ (because $c \in C$ implies $1 \in \{ t > 0 : c \in t C \}$). Next $A \cap B = \emptyset$ implies $x_0 \notin C$ and $p(x_0) \geq 1$.

Let $M$ be the linear span of $\{ x_0 \}$. Define $f$ on $M$ by $f(tx_0) = t$. Then $f(x_0) = 1 \leq p(x_0)$ implies $f(x) \leq p(x)$ on $M$. So $f$ can be extended linearly to $X$ with $f(x) \leq p(x)$ on $X$. Since $f(x) \leq p(x) \leq 1$ for all $x \in C$, so $f(-x) = -f(x) \geq -1$ for all $-x \in -C$. Then $|f| \leq 1$ on $U = C \cap (-C)$, a neighborhood of 0. Thus, for all $\epsilon > 0$, $\epsilon U$ is a neighborhood of 0 and $x \in \epsilon U$ implies $|f(x)| \leq \epsilon$. So $f$ is continuous at 0, hence continuous on $X$.

For all $a \in A$ and $b \in B$, $f(a) - f(b) + 1 = f(a - b + x_0) \leq 1$ so that $f(a) \leq f(b)$. Since $A$ is nonempty open convex, $f \in X^*$ and $f \neq 0$, by the lemma, $f(A) = (\alpha, \beta)$ say. Then for all $a \in A$, we have $f(a) < \beta \leq \inf \{ f(b) : b \in B \}$.

(b) Since $A \cap B = \emptyset$, $B$ is closed and $X$ is locally convex, $X \setminus B$ is a neighborhood of every $a \in A$. So there is an open convex neighborhood $V_0$ of 0 such that $a + V_0 \subseteq X \setminus B$. Now $\{ a + \frac{1}{2} V_0 \}$ covers $A$. From a subcover $\{ a_i + \frac{1}{2} V_{a_i} : i = 1, 2, \ldots, n \}$, we intersect the $\frac{1}{2} V_{a_i}$’s to get an open convex neighborhood $V$ of 0. Note

$$A + V \subseteq \bigcup_{i=1}^{n} (a_i + \frac{1}{2} V_{a_i} + V) \subseteq \bigcup_{i=1}^{n} (a_i + \frac{1}{2} V_{a_i} + \frac{1}{2} V_{a_i}) \subseteq \bigcup_{i=1}^{n} (a_i + V_{a_i}) \subseteq X \setminus B.$$
Then $A + V$ is an open convex set disjoint from $B$. By (a), there is a continuous linear functional $f : X \to \mathbb{R}$ such that $f < \beta \leq \inf \{ f(y) : y \in B \}$ on $A + V$ and $f \geq \beta$ on $B$. Since $f(A)$ is compact in $(-\infty, \beta)$, we have

$$\sup \{ f(x) : x \in A \} < \beta \leq \inf \{ f(y) : y \in B \}. \quad \square$$

**Remark.** (b) is often called the Strong Separation Theorem.

**Corollary (Consequences of Separation Theorem).** Let $X$ be a locally convex space.

(a) If $X$ is Hausdorff, then $X^*$ separates points of $X$ in the sense that for every $x, y \in X$ with $x \neq y$, there exists $f \in X^*$ such that $f(x) \neq f(y)$. In particular, if $f(x) = 0$ for all $f \in X^*$, then $x = 0$.

(b) Let $Y$ be a vector subspace of $X$ and $x \in X$. We have $x \notin \overline{Y}$ if and only if there exists $f \in X^*$ such that $f(x) \neq 0$ and $f \equiv 0$ on $Y$. Also, $\overline{Y} = X$ if and only if $f \in X^*$ with $f \equiv 0$ on $Y$ implies $f \equiv 0$ on $X$.

**Proof.** (a) For distinct $x, y \in X$, let $A = \{x\}$ and $B = \{y\}$ and apply (b) of the separation theorem.

(b) For the if direction, by continuity, $f \equiv 0$ on $\overline{Y}$ and so $x \notin \overline{Y}$. For the only-if direction, let $A = \{x\}$ and $B = \overline{Y}$ and apply (b) of the separation theorem to get $f \in X^*$ to separate $A$ and $B$. Since $f(Y)$ is a vector subspace of $\mathbb{R}$, we must have $f(Y) = \{0\}$ and $f(x) \neq 0$.

Using the separation theorem, we can obtain an important theorem of M. Krein and D. Milman.

**Definitions.** Let $S$ be a nonempty subset of a vector space $V$ over $\mathbb{K}$.

(a) A nonempty subset $M$ of $S$ is an **extreme set** in $S$ iff $M$ has the property that “if there exist $s_1, s_2 \in S$ and there exists $t \in (0, 1)$ such that $ts_1 + (1-t)s_2$ is in $M$, then both $s_1$ and $s_2$ are in $M$.” An extremal set consisted of a single point is called an **extreme point**.

(b) The **convex hull** of $S$ is the smallest convex set in $V$ containing $S$. (It is easy to see that the convex hull of $S$ is $\{ \sum_{i=1}^{n} t_i s_i : s_i \in S, t_i \in [0, 1], \sum_{i=1}^{n} t_i = 1 \}$.) For $S$ in a topological vector space, the **closed convex hull** of $S$ is the closure of the convex hull of $S$.

**Examples.** The sides of a triangular region on a plane are extreme sets of the region and the vertices are extreme points. Every point of a circle is an extreme point of the closed disk having the circle as boundary.

**Remarks.** (1) If for every $\alpha \in A$, $E_\alpha$ is an extreme set in $S$ and $E = \bigcap_{\alpha \in A} E_\alpha \neq \emptyset$, then $E$ is an extreme set in $S$. This is because $s_1, s_2 \in S$, $t \in (0, 1)$ and $ts_1 + (1-t)s_2 \in E$ imply $ts_1 + (1-t)s_2 \in E_\alpha$ for every $\alpha \in A$, which implies $s_1, s_2 \in E_\alpha$ for every $\alpha$, hence $s_1, s_2 \in E$.

(2) If $P$ is an extreme set in $H_1$ and $H_1$ is an extreme set in $S$, then $P$ is an extreme set in $S$. This is because $ts_1 + (1-t)s_2 \in P$ for some $s_1, s_2 \in S$, $0 < t < 1$ implies $ts_1 + (1-t)s_2 \in H_1$ so that $s_1, s_2 \in H_1$ (by the extremity of $H_1$ in $S$), then $s_1, s_2 \in P$ (by the extremity of $P$ in $H_1$).

**Theorem (Krein-Milman).** Let $X$ be a Hausdorff locally convex space and $\emptyset \neq S \subseteq X$. If $S$ is compact and convex, then $S$ has at least one extreme point and $S$ is the closed convex hull of its extreme points.

**Proof.** We first show $S$ has an extreme point. Note $S$ is an extreme subset of itself. Let $\mathcal{C}$ be the collection of all nonempty compact extreme subsets of $S$. Order $\mathcal{C}$ by reverse inclusion, i.e. for $E_1, E_2 \in \mathcal{C}$, define $E_1 \preceq E_2$ iff $E_1 \supseteq E_2$. For every nonempty chain in $\mathcal{C}$, since $X$ is Hausdorff, elements of the chain are closed. By the finite intersection property, the intersection of all elements of the chain is nonempty and closed (hence compact). By remark (1), it is an extreme subset of $S$. So it is an upper bound of the chain in $\mathcal{C}$.

By Zorn’s lemma, $\mathcal{C}$ has a maximal element $E$. Assume $E$ has distinct elements $x, y$. By (b) of the separation theorem, there exists $f \in X^*$ such that $\text{Re } f(x) < \text{Re } f(y)$. This implies $y \notin E_0 = \{ s \in E : \text{Re } f(s) = \inf \text{Re } f(E) \} \subset E$. Now $E_0$ is nonempty due to continuity of $\text{Re } f$ on the compact set $E$. Since
The Krein-Milman theorem can be used to prove the Stone-Weierstrass theorem. Combining of
For every
\( E_0 = (\operatorname{Re} f)^{-1}(\{\inf \operatorname{Re} f(E)\}) \), it is closed (hence compact). Finally, \( E_0 \) is an extreme subset of \( S \) because \( s = t s_1 + (1 - t) s_2 \in E_0 \subseteq E \) implies \( s_1, s_2 \in E \) (as \( E \) is extreme) and
\[
\inf \operatorname{Re} f(E) \leq \min\{\operatorname{Re} f(s_1), \operatorname{Re} f(s_2)\} \leq t \operatorname{Re} f(s_1) + (1 - t) \operatorname{Re} f(s_2) = \operatorname{Re} f(s) = \inf \operatorname{Re} f(E)
\]
implies \( \operatorname{Re} f(s_1) = \inf \operatorname{Re} f(E) = \operatorname{Re} f(s_2) \), i.e. \( s_1, s_2 \in E_0 \). Since \( E_0 \supseteq E \), this contradicts the maximality of \( E \) in \( C \). Therefore, \( E \) can only be an extreme point of \( S \).

Now we show \( S \) equals the closed convex hull \( H \) of its extreme points. Since \( S \) is closed and convex, \( H \subseteq S \). Assume there is \( s \in S \setminus H \). By (b) of the separation theorem, there is \( f \in X^* \) such that \( \operatorname{Re} f(s) < \inf \{ \operatorname{Re} f(y) : y \in H \} \). Then \( H_1 = \{ x \in S : \operatorname{Re} f(x) = \inf \operatorname{Re} f(S) \} \) is convex and disjoint from \( H \). Similar to \( E_0 \) above, \( H_1 \) is a nonempty closed (hence compact) extreme subset of \( S \). By the first part, \( H_1 \) has at least one extreme point \( p \). By remark (2), \( p \) is an extreme point of \( S \), which contradicts \( H_1 \cap H = \emptyset \). So \( S = H \) \( \square \)

**Remarks.** The Krein-Milman theorem can be used to prove the Stone-Weierstrass theorem. Combining with the Banach-Alaoğlu theorem in the next chapter, it can be used to show that there exist Banach spaces that are not the dual spaces of Banach spaces. For details of these two applications, see [Be], p. 110.

---

**Applications of Theorems in Chapter 2.**

We first remark that every function \( f \) defined on \( (-\pi, \pi] \) corresponds to a \( 2\pi \)-periodic function on \( \mathbb{R} \) defined by \( f(x + 2n\pi) = f(x) \) for all integers \( n \). Let \( e^{i\theta} = \cos \theta + i \sin \theta \) and \( T = \{ e^{i\theta} : -\pi < \theta \leq \pi \} \).

Every function \( f \) defined on \( (-\pi, \pi] \) also corresponds to a function \( f_0 \) on \( T \) defined by \( f_0(e^{i\theta}) = f(\theta) \).

**In the following we will use these correspondences to identify these three sets of functions.**

**Definitions.** (1) A function \( P : \mathbb{R} \to \mathbb{C} \) is a trigonometric polynomial iff it is of the form \( P(x) = \sum_{k=-n}^{n} c_k e^{ikx} \), where \( c_k \in \mathbb{C} \) and \( n \) is a nonnegative integer.

(2) For all \( f \in L^1(-\pi, \pi] \) and \( n \in \mathbb{Z} \), define the \( n \)-th Fourier coefficient of \( f \) to be \( \hat{f}(n) = \int_{[-\pi,\pi]} f(x) e^{-inx} \, dm \). The Fourier series of \( f \) is \( \sum_{k=-\infty}^{\infty} \hat{f}(k) e^{ikx} \) and its \( n \)-th partial sum is \( s_n(f; x) = \sum_{k=-n}^{n} \hat{f}(k) e^{ikx} \).

**Remarks.** (1) Under the identification above, since the trigonometric polynomials are \( 2\pi \)-periodic on \( \mathbb{R} \), they can be considered as functions on \( T \). Below \( 2\pi \)-periodic continuous functions on \( \mathbb{R} \) will be considered as functions in \( C(T) \). Functions in \( L^1(-\pi, \pi] \) can be considered as functions in \( L^1(T) \).

(2) The set of all trigonometric polynomials is dense in \( C(T) \) with sup-norm by the Stone-Weierstrass theorem since it is a self-adjoint subalgebra of \( C(T) \) that separates points of \( T \) and vanishes at no point of \( T \).

(3) The Dirichlet kernel is \( D_n(x) = \sum_{k=-n}^{n} e^{ikx} \), which is \( \frac{\sin(n + \frac{1}{2})x}{\sin \frac{1}{2}x} \) if \( x \neq 0 \) and is \( 2n + 1 \) if \( x = 0 \). We have
\[
s_n(f; x) = \sum_{k=-n}^{n} \hat{f}(k) e^{ikx} = \sum_{k=-n}^{n} \int_{[-\pi,\pi]} f(\theta) e^{ik(x-\theta)} \, dm \frac{d\theta}{2\pi} = \int_{[-\pi,\pi]} f(\theta) D_n(x-\theta) \, d\theta = (f * D_n)(x).
\]

**Riemann-Lebesgue Lemma.** For every \( f \in L^1(T) \), \( \lim_{n \to \pm \infty} \hat{f}(n) = 0 \). In fact, the function \( \mathcal{F} : L^1(T) \to c_0 \) defined by \( \mathcal{F}(f) = (\hat{f}(0), \hat{f}(1), \hat{f}(-1), \hat{f}(2), \hat{f}(-2), \ldots) \) is continuous and linear.
Proof. For every $\varepsilon > 0$, from measure theory (see Rudin, *Real and Complex Analysis*, Theorem 3.14), there exists $g \in C(\mathbb{T})$ such that $\|f - g\|_\infty < \varepsilon/2$. Next by remark (2) above, there is a trigonometric polynomial

$$P(x) = \sum_{k=-N}^{N} c_k e^{ikx}$$

such that $\|g - P\|_\infty < \varepsilon/2$. For $|n| > N$, we have $\hat{P}(n) = 0$ and

$$|\hat{f}(n)| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(t) - P(t)) e^{-int} dt \right| \leq \|f - P\|_1 \leq \|f - g\|_1 + \|g - P\|_1 \leq \|f - g\|_1 + \|g\|_\infty < \varepsilon.$$

So $\hat{f}(n) \to 0$ as $|n| \to \infty$.

Next, linearity of $F$ is clear and continuity follows from $\|F(f)\| = \sup |\hat{f}(n)| \leq \int_{(-\pi,\pi]} |f| dm = \|f\|_1$.

**Questions** Is $F$ injective? Is it surjective?

**Theorem.** $F : L^1(\mathbb{T}) \to c_0$ is injective.

**Proof.** Suppose $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$. Then $\int_{(-\pi,\pi]} fP dm = 0$ for all trigonometric polynomials $P$. There are two ways to finish.

1) By remark (2) above, we have $\int_{(-\pi,\pi]} fg dm = 0$ for all $g \in C(\mathbb{T})$. For those who know the Riesz representation theorem on $C(\mathbb{T}^*)$, it follows $f = 0$ almost everywhere.

2) For every $x \in (-\pi,\pi]$, there are continuous $g_n : (-\pi,\pi] \to [0,1]$ such that $g_n(-\pi) = g_n(\pi) = 0$ and $\lim_{n \to \infty} g_n(t) = \chi(-\pi,x)(t)$ for all $t \in (-\pi,\pi]$. By remark (2) above, there is a trigonometric polynomial $P_n$ such that $\|g_n - P_n\|_\infty < \frac{1}{n}$. Then $\|fP_n\|_1 \leq \|f\|_1 (\|g_n\|_\infty + \frac{1}{n}) \leq 2\|f\|_1$ and $f(t)P_n(t) \to f(t)\chi(-\pi,x)(t)$ for all $t \in (-\pi,\pi]$. By the Lebesgue dominated convergence theorem,

$$\int_{-\pi}^{\pi} f(t) dt = \int_{(-\pi,\pi]} f\chi(-\pi,x) dm = \lim_{n \to \infty} \int_{(-\pi,\pi]} fP_n dm = 0.$$

Differentiate with respect to $x$, we get $f = 0$ almost everywhere (see Rudin, *Real and Complex Analysis*, Theorem 7.11). \qed

**Theorem.** $F : L^1(\mathbb{T}) \to c_0$ is not surjective. In fact, the range of $F$ is not closed.

**Proof.** Assume $F$ is surjective. There are two ways to get a contradiction.

1) By the inverse mapping theorem, $F$ would be an isomorphism between $L^1(\mathbb{T})$ and $c_0$. Then $c_0^* = \ell^1$ would be isomorphic to $(L^1(\mathbb{T}))^* = L^\infty(\mathbb{T})$, which is impossible because $\ell^1$ is separable, but $L^\infty(\mathbb{T})$ (like $\ell^\infty$) is not separable as there are uncountably many balls $\{B(\chi(-\pi,x), \frac{1}{n}) : x \in (-\pi,\pi]\}$ that are pairwise disjoint in $L^\infty(\mathbb{T})$. Hence, we have a contradiction.

2) Since $F$ is injective, if $F(L^1(\mathbb{T})) = c_0$ or closed, then by the lower bound theorem, $F$ would be bounded below, i.e. there exists $c > 0$ such that $\|F(f)\|_\infty \geq c\|f\|_1$ for all $f \in L^1(\mathbb{T})$. Now $D_n \in C(\mathbb{T}) \subseteq L^1(\mathbb{T})$ and $\|F(D_n)\|_\infty = \|[1,1,\ldots,1,0,0,\ldots]\|_\infty = 1$. However, since $|\sin x| \leq |x|$ for all $x \in \mathbb{R}$, we have

$$\|D_n\|_1 > \frac{2}{\pi} \int_0^{\pi} |\sin(n + \frac{1}{2})\theta| d\theta = \frac{2}{\pi} \int_0^{(n+1/2)\pi} |\sin\phi| d\phi > \frac{2}{\pi} \sum_{k=1}^{n} \frac{1}{k} \int_{(k-1)\pi}^{k\pi} |\sin\phi| d\phi = \frac{4}{\pi} \sum_{k=1}^{n} \frac{1}{k} \to \infty,$$

which contradicts $F$ is bounded below. \qed

**Questions** Does the Fourier series of $f \in L^1(\mathbb{T})$ converge to $f$ almost everywhere or in $L^1$-norm?

**Theorem (du Bois-Reymond, 1873).** For every $w \in (-\pi,\pi]$, there exists $f \in C(\mathbb{T})$ such that its Fourier series diverges at $x = w$. More precisely, the partial sums of the Fourier series at $x = w$ is unbounded.
Proof. (Due to Henri Lebesgue) First we deal with the case \( w = 0 \). Define \( T_n : C(\mathbb{T}) \rightarrow \mathbb{C} \) by \( T_n(f) = s_n(f;0) = \sum_{k=-n}^{n} \hat{f}(n) \). Clearly, \( T_n \) is linear. Also, \( T_n \) is bounded since

\[
|T_n f| = \left| \int_{-\pi}^{\pi} f(\theta)D_n(\theta) \frac{dm}{2\pi} \right| \leq \|f\|_\infty \int_{-\pi}^{\pi} |D_n(\theta)| d\theta = \|D_n\|_1 \|f\|_\infty.
\]

So \( \|T_n\| \leq \|D_n\|_1 \).

In fact, \( \|T_n\| = \|D_n\|_1 \). To see this, let \( g(t) = \text{sgn} D_n(-t) \), which is defined by \( g(t) = 1 \) if \( D_n(-t) \geq 0 \) and \( g(t) = -1 \) if \( D_n(-t) < 0 \). Then \( g(t)D_n(-t) = \|D_n(-t)\| \). Also, there exists \( f_j \in C(\mathbb{T}) \) such that \( \|f_j\|_\infty = 1 \) and \( \lim_{j \to \infty} f_j(t) = g(t) \) for every \( t \in (-\pi,\pi] \). Since \( f_j(\theta)D_n(\theta) \rightarrow g(\theta)D_n(\theta) = \|D_n(-\theta)\| \) and \( |f_j(\theta)D_n(\theta)| \leq \|D_n(-\theta)\| \in C(\mathbb{T}) \subset L^1(\mathbb{T}) \), by the Lebesgue dominated convergence theorem,

\[
\lim_{j \to \infty} T_n f_j = \lim_{j \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} f_j(\theta)D_n(\theta) d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\theta)D_n(\theta) d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(\theta)| d\theta = \|D_n\|_1.
\]

Now \( \sup_{n \geq 0} \|T_n\| : n = 0,1,2,\ldots \) is bounded, hence \( \sup_{n \geq 0} \|T_n\| = \|D_n\|_1 = \infty \). By the uniform boundedness principle, there exists \( f \in C(\mathbb{T}) \) such that \( \sup_{n \geq 0} \|T_n f\| = \|f\|_\infty \). Therefore, the Fourier series of \( f \) converges when \( w = 0 \). For \( w \neq 0 \), \( f_w(x) = f(x-w) \in C(\mathbb{T}) \) has Fourier coefficients \( \hat{f}_w(k) = \hat{f}(k)e^{-ikw} \). Hence, its Fourier series is

\[
\sum_{k=-\infty}^{\infty} (\hat{f}(k)e^{-ikw})e^{ikx},
\]

which diverges at \( x = w \).

Principle of Condensation of Singularities. Let \( X \) be a Banach space and \( Y \) be a normed space. Let \( T_{nj} \in L(X,Y) \) for \( n,j = 0,1,2,\ldots \) be such that for all \( j \), \( \limsup_{n} \|T_{nj}\| = \infty \). Then there is a set \( U \) of second category in \( X \) such that for all \( f \in U \) and all \( j \), \( \limsup_{n} \|T_{nj} f\| = \infty \).

Proof. For a fixed \( j \), let \( V_j = \{ f \in X : \limsup_{n} \|T_{nj} f\| < \infty \} \). Then \( f \in V_j \) implies \( \sup_{n \geq 0} \|T_{nj} f\| : n = 0,1,2,\ldots \) is bounded, hence \( \limsup_{n} \|T_{nj}\| < \infty \), a contradiction. So \( V_j \) is of first category in \( X \). Then \( V = \bigcup_{j \geq 0} V_j \) is of first category in \( X \). Since \( X \) is complete, \( U = X \setminus V \) is of second category in \( X \). For all \( f \in U \) and all \( j \), we have \( f \notin V_j \), i.e. \( \limsup_{n} \|T_{nj} f\| = \infty \).

Application. Now take a countable dense subset \( \{w_j\} \) of \( \mathbb{T} \) and define \( T_{nj} : C(\mathbb{T}) \rightarrow \mathbb{C} \) by \( T_{nj} f = s_n(f;w_j) \). As in the proof of the last theorem, \( \|T_{nj}\| = \|D_n\|_1 \) and \( \limsup_{n \to \infty} \|T_{nj}\| = \infty \). By the principle of condensation of singularities, there is a set of second category in \( C(\mathbb{T}) \) such that all these functions have Fourier series diverging at the dense subset \( \{w_j\} \) (with (*) \( \sup\{s_n(f,w_j) : n = 1,2,3,\ldots \} = \infty \) for all \( w_j \)).

Let \( f \) be one such function. We claim that the set of points on \( \mathbb{T} \) where the Fourier series of \( f \) diverges is actually a set of second category in \( \mathbb{T} \), hence uncountable and much more than \( \{w_j\}! \).

To see this, let \( M_{n,k} = \{ w \in \mathbb{T} : |s_n(f,w)| \leq k \} \), \( M_k = \bigcap_{n=1}^{\infty} M_{n,k} \) and \( M = \bigcup_{k=1}^{\infty} M_k \).

1. \( M_k = \{ w \in \mathbb{T} : \sup\{ |s_n(f,w)| : n = 1,2,3,\ldots \} \leq k \} \), so by (*), for all \( j,k \), \( w_j \notin M_k \).

2. If the Fourier series of \( f \) converges at \( w \), then \( s_n(f,w) : n = 1,2,3,\ldots \) is bounded, hence \( w \) is in some \( M_k \), leading to \( w \in M \). In particular, the Fourier series of \( f \) diverges at all elements of \( \mathbb{T} \setminus M \).

3. \( h_{n,j}(w) = s_n(f,w) = (f * D_n)(w) \) is continuous in \( w \). So \( M_{n,k} = h_{n,j}^{-1}(B(0,k)) \) and \( M_k \) are closed.
Assume some $M_k$ is of second category in $T$. Then in particular, it would not be nowhere dense. Since $M_k$ is closed by (3), there is a nonempty open set in $M_k$. By the density of $\{w_j\}$, one of the $w_j$ would be in $M_k$, contradicting (1). So all $M_k$ must be of first category in $T$. Then $M$ will also be of first category in $T$. By (2), the Fourier series of $f$ diverges on $T \setminus M$, which is of second category in $T$, hence uncountable!

**Remarks.** In 1915, Lusin conjectured that for $f \in L^2(-\pi, \pi)$, the Fourier series of $f$ converges almost everywhere.

In 1926, Kolmogorov (as an undergraduate student in Moscow State University) proved that there exists a $f \in L^1(-\pi, \pi]$ such that the Fourier series of $f$ diverges everywhere! See Antoni Zygmund, *Trigonometric Series*, second edition, vol. 1, pp. 310-314 for such a function.

In 1927, M. Riesz proved that for every function $f$ in $L^p(-\pi, \pi]$ ($1 < p < \infty$), the Fourier series of $f$ converges in the $L^p$-norm to $f$. From measure theory (see Rudin, *Real and Complex Analysis*, Theorem 3.12), it is known that this implies there is a subsequence of the partial sums of the Fourier series of $f \in L^p(-\pi, \pi]$ converging almost everywhere to $f$.

In 1966, Lennart Carleson proved the Lusin conjecture. In particular, this implies the Fourier series of $2\pi$-periodic continuous functions converge almost everywhere (to itself by Riesz’ result). In the same year, Kahane and Katznelson proved that for every set of Lebesgue measure 0 on $(-\pi, \pi]$, there is a $2\pi$-periodic continuous function whose Fourier series diverges there.

In 1968, Richard Hunt proved that for every $f \in L^p(-\pi, \pi]$ with $1 < p \leq \infty$, the Fourier series of $f$ converges almost everywhere to itself.
Chapter 3. Weak Topologies and Reflexivity.

§1. Canonical Embedding. For a normed space $X$ over $\mathbb{K}$, $x \in X$ and $y \in X^*$, let $\langle x, y \rangle = y(x)$. This notation is to illustrate that many similar properties exist between $X$ and $X^*$. For example, $(x, y)$ is linear in $x$ and $y$. For $y \in X^*$, $\|y\| = \sup\{|y(x)| : x \in X, \|x\| \leq 1\} = \sup\{|\langle x, y \rangle| : x \in X, \|x\| \leq 1\}$. In remark (1) below, we will show that $\|x\| = \sup\{|y(x)| : y \in X^*, \|y\| \leq 1\} = \sup\{|\langle x, y \rangle| : y \in X^*, \|y\| \leq 1\}$.

**Theorem.** Let $X, Y$ be normed spaces. If $Y$ is complete, then $L(X,Y)$ is a Banach space. (In particular, $X^* = L(X, \mathbb{K})$ is a Banach space.)

**Proof.** Clearly $L(X,Y)$ is a normed vector space. For completeness, suppose $\{T_n\}$ is a Cauchy sequence in $L(X,Y)$. Then $\{T_n\}$ is bounded so that there is $K \geq 0$ such that for all $n \geq 1$, $\|T_n\| \leq K$. Then for all $x \in X$, $n \geq 1$, we have $\|T_n(x)\| \leq K \|x\|$. Since $\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \|x\|$, the sequence $\{T_n(x)\}$ is a Cauchy sequence in $Y$. Since $Y$ is complete, $\lim_{n \to \infty} T_n(x)$ exists and we may define $T(x) = \lim_{n \to \infty} T_n(x)$. Clearly, $T$ is linear. Also, $T$ is bounded as $\|T(x)\| = \lim_{n \to \infty} \|T_n(x)\| \leq K \|x\|$. So $T \in L(X,Y)$.

For every $\varepsilon > 0$, since $\{T_n\}$ is Cauchy, there is such that $n, m \geq N$ implies $\|T_n - T_m\| < \varepsilon$. Then $\|T_n(x) - T_m(x)\| \leq \varepsilon \|x\|$ for all $x \in X$. So $\|T_n(x) - T(x)\| = \lim_{m \to \infty} \|T_n(x) - T_m(x)\| \leq \varepsilon \|x\|$. Hence, if $n \geq N$, then $\|T_n - T\| \leq \varepsilon$. Therefore, $\{T_n\}$ converges to $T$ in $L(X,Y)$. \[\Box\]

**Exercise.** For $X \neq \{0\}$, if $L(X,Y)$ is a Banach space, then prove that $Y$ is complete.

**Canonical Embedding Theorem.** For a normed space $X$, the “canonical embedding” $i : X \to X^{**} = (X^*)^*$ defined by $i(x) = i_x$, where $i_x(y) = y(x)$, is a linear isometry. If $X \neq \{0\}$, then for all $x \in X$, $\|x\| = \sup\{|y(x)| : y \in X^*, \|y\| = 1\}$.

**Proof.** It is easy to see that $i_x$ is a linear transformation from $X^*$ to $\mathbb{K}$ and $i$ is a linear operator from $X$ to $X^{**}$. To show $i$ is an isometry, it is enough to deal with the case $X \neq \{0\}$. Note $|i_x(y)| = |y(x)| \leq \|y\| \|x\|$ for all $y \in X^*$ so that $\|i_x\| \leq \|x\|$. By part (c) of the Hahn-Banach theorem, for every $x \in X$, there is $y \in X^*$ such that $\|y\| = 1$ and $y(x) = \|x\|$. Then $\|x\| = y(x) = i_x(y) \leq \|i_x\| \|y\| = \|i_x\|$. Therefore, $\|x\| = \|i_x\|$.

**Remarks.** (1) In the case $X = \{0\}$, we have $X^* = \{0\}$. So to cover all normed spaces, the second statement should be changed to $\|x\| = \sup\{|y(x)| : y \in X^*, \|y\| \leq 1\} = \sup\{|\langle x, y \rangle| : y \in X^*, \|y\| \leq 1\}$.

(2) When notations become cumbersome, we will identify $x \in X$ with $i_x \in X^{**}$. Also, we will often identify $X$ with $i(X)$ below.

**Definitions.** The closure $\hat{X}$ of $X$ in $X^{**}$ is a Banach space containing $X$ as a dense subset and it is called a completion of $X$. Banach spaces $X$ satisfying $i(X) = X^{**}$ are called reflexive. (For example, Hilbert spaces, $L^p([0,1])$ and $\ell^p$ with $1 < p < \infty$ are reflexive.)

§2. Weak Topologies. For a normed space $X$, there is a weakest vector topology $w$ on $X$ that makes all elements of $X^*$ continuous. We simply take $\mathcal{P} = \{\{f\} : f \in X^*\}$ and apply the theorems on locally convex spaces. This topology $w$ on $X$ is called the weak topology on $X$. Then $X$ with this topology is a locally convex space. Using the description of a base of $0$ in a locally convex space, we see sets of the form $U = \bigcap_{i=1}^{n} \{x \in X : |f_i(x)| < r_i\}$, where $r_i > 0$ and $f_i \in X^*$, form a base at $0$ for the weak topology.

So on a normed space $X$, there are two topologies, namely the original norm-topology and the $w$-topology. When we mean $X$ with the $w$-topology, we shall write $(X, w)$.

**Properties of Weak Topologies.**

(1) By definition of weak topology, we have $w$-topology is a subset of the norm-topology. So $w$-open sets are open in $X$, $w$-closed sets are closed in $X$, but compact sets in $X$ are $w$-compact.
(2) By part (a) of the corollary following the separation theorem, $X^*$ separates points of $X$, which implies the weak topology is Hausdorff. So $w$-compact sets are $w$-closed.

(3) For every net $\{x_\alpha\}_{\alpha \in I}$ in $X$, by a theorem in the section on locally convex spaces, we have $\{x_\alpha\}_{\alpha \in I}$ $w$-converges to $x$ in $X$ (write as $x_\alpha \rightharpoonup^w x$) iff for every $f \in X^*$, $|f(x_\alpha - x)| \to 0$, i.e. $f(x_\alpha) \to f(x)$.

(4) For a normed space $X$, a sequence $x_n \rightharpoonup^w x$ in $X$ iff there is $C > 0$ such that $\|x_n\| < C$ for $n = 1, 2, 3, \ldots$ and $M = \{f \in X^* : \lim_{n \to \infty} f(x_n) = f(x)\}$ is dense in $X^*$. This follows from the uniform boundedness principle, part (b) of the Banach-Steinhaus theorem and the canonical embedding theorem that $\|x_n\| = \|f(x_n)\|$.

(5) For a convex subset $C$ of a normed space $X$, we have $\overline{C} = \overline{C}^w$. $C$ is closed iff it is $w$-closed. $C$ is dense iff it is $w$-dense.

**Proof.** Conversely, assume there is $x_0 \in \overline{C}^w \setminus \overline{C}$. By the separation theorem, there is $f \in X^*$ such that $\Re f(x_0) < s = \inf\{\Re f(x) : x \in \overline{C}\}$. Since $f$ is $w$-continuous, $U = \{x \in X : \Re f(x) < s\} = f^{-1}(\{z \in K : \Re z < s\})$ is a $w$-open neighborhood of $x_0$ and disjoint from $C$, hence also from $\overline{C}^w$. So $x_0 \notin \overline{C}^w$, a contradiction. Therefore $\overline{C} = \overline{C}^w$. The second and third statements follow easily from the first statement.

§3. **Weak-star Topologies.** Similarly, on a dual space $X^* = L(X, K)$ (which is a normed space), for each $x \in X$, consider $i_x$ in the canonical embedding. We can take $P = \{i_x : x \in X\}$ to generate a topology $w^*$ on $X^*$ so that all $i_x$ are continuous. This topology $w^*$ on $X^*$ is called the *weak-star topology* on $X^*$. Then $X^*$ with this topology is a locally convex space. Using the description of a base of $0$ in a locally convex space, we see sets of the form $U^* = \bigcap_{i=1}^n \{f \in X^* : |f(x_i)| < r_i\}$, where $r_i > 0$ and $x_i \in X$, form a base at $0$ for the weak-star topology.

Thus, on a dual space $X^*$, there are more than one topologies we will be using, namely the original norm-topology and the $w^*$-topology. When we mean $X$ with $w^*$-topology, we shall write $(X^*, w^*)$.

**Properties of Weak-star Topologies.**

(1) By definition of weak-star topology, we have the $w^*$-topology is a subset of the norm-topology. So $w^*$-open sets are open in $X^*$, $w^*$-closed sets are closed in $X^*$, but compact sets in $X$ are $w^*$-compact.

(2) $i(X)$ separates points of $X^*$. This implies the $w^*$ topology is Hausdorff and $w^*$-compact sets are $w^*$-closed.

(3) For a net $\{f_\alpha\}_{\alpha \in J}$ in $X^*$, we have $\{f_\beta\}_{\beta \in J}$ $w^*$-converges to $f$ in $X^*$ (write as $f_\beta \rightharpoonup^w f$) iff for every $x \in X$, $f_\beta(x) \to f(x)$.

(4) Let $X$ be a Banach space. A sequence $f_n \rightharpoonup^w f$ in $X^*$ iff there is $C > 0$ such that $\|f_n\| < C$ for $n = 1, 2, 3, \ldots$ and $M = \{x \in X : \lim_{n \to \infty} f_n(x) = f(x)\}$ is dense in $X$. This follows from the uniform boundedness principle and part (b) of the Banach-Steinhaus theorem.

Next, we will show that for a convex subset $C$ of a dual space $X^*$, $\overline{C} = \overline{C}^w$ may not hold.

**Lemma.** Let $g, g_1, \ldots, g_n$ be linear functionals on a vector space $X$. If $\bigcap_{i=1}^n \ker g_i \subseteq \ker g$, then there are $c_1, \ldots, c_n \in K$ such that $g = c_1 g_1 + \cdots + c_n g_n$. The converse is trivially true.

**Proof.** Define $T : X \to K^n$ by $T(x) = (g_1(x), \ldots, g_n(x))$. Then $\ker T = \ker g_1 \cap \cdots \cap \ker g_n$. If $T(x) = T(x')$, then $x - x' \in \ker T \subseteq \ker g$ and so $g(x) = g(x')$. Choose a basis for $\text{ran} T$ and extend it to a basis for $K^n$. Define $G : K^n \to K$ by $G(T(x)) = g(x)$ for $x \in X$ and $G(v) = 0$ for $v$ in the extended part of the basis. Then $G$ is linear and $g = G \circ T$. For the standard basis $\{e_1, \ldots, e_n\}$ of $K^n$, let $c_1 = G(e_1)$, then $G(x_1, \ldots, x_n) = G(x_1 e_1 + \cdots + x_n e_n) = c_1 x_1 + \cdots + c_n x_n$. Therefore, $g = G \circ T = c_1 g_1 + \cdots + c_n g_n$. 

**Weak-star Functional Theorem.** Let $X$ be a normed space. If $g : X^* \to K$ is linear and continuous with the weak-star topology on $X^*$, then $g = i_x$ for some $x \in X$.  

25
Proof. Since $g^{-1}\{\{c \in K : |c| < 1\}\}$ is a $w^*$-open set containing 0, there is a $w^*$-neighborhood $\{z \in X^* : |z(x_j)| < r_1, \ldots, |z(x_n)| < r_n\}$ of 0 in $g^{-1}\{\{c \in K : |c| < 1\}\}$. Now for fixed $r > 0$, $i_x(r) = g(x) = 0$ if and only if for all $t > 0$, $|g(x)| < tr$ if and only if for all $t > 0$, $y = tz$ with $|z(x)| < r$. Using this, we have

$$\bigcap_{j=1}^n \ker i_{x_j} = \bigcap_{t>0} \bigcap_{j=1}^n r\{z \in X^* : |z(x_j)| < r_j\} = \bigcap_{t>0} \bigcap_{j=1}^n r\{z \in X^* : |z(x_j)| < r_j\} \subseteq \bigcap_{t>0} g^{-1}\{\{c \in K : |c| < 1\}\} = \ker g.$$

By the last lemma, this implies $g = c_1i_{x_1} + \cdots + c_ni_{x_n} = i_x$, where $x = c_1x_1 + \cdots + c_nx_n$. □

Remark. Now we show for a convex subset $C$ of a dual space $X^*$, $\overline{C} = \overline{C}^w$ may not hold. Let $X$ be an incomplete normed space. Take a $g \in X^* \setminus i(X)$, then $C = \ker g$ is convex and norm-closed in $X^*$. If $C = \ker g$ is $w^*$-closed, then by the closed kernel theorem, $g$ would be a $w^*$-continuous linear functional, hence in $i(X)$ by the weak-star functional theorem, a contradiction.

Lemma. In a topological space $X$, a point $x$ is in the closure of a set $S$ if and only if every (open) neighborhood of $x$ intersects $S$.

Proof. Equivalently, we need to show $x \not\in \overline{S}$ iff there is an open neighborhood of $x$ disjoint from $S$. So if $x \not\in \overline{S}$, then $x \in X \setminus \overline{S}$, which is open and disjoint from $S$. Conversely, if there is an open neighborhood $U$ of $x$ disjoint from $S$, then $x \not\in X \setminus U$, which is closed and contains $S$. Hence, $X \setminus U \subseteq \overline{S}$. So $x \not\in \overline{S}$. □

Theorem (Tychonoff). The Cartesian product $S$ of a family of compact spaces $\{S_\alpha : \alpha \in A\}$ is compact.

Proof. (Due to Nicholas Bourbaki) Let $\mathcal{F}$ be any collection of closed subsets of $S$ having the finite intersection property. We need to prove $\bigcap \mathcal{F} \neq \emptyset$.

Let $P(S)$ be the set of all subsets of $S$ and $\Omega = \{A : \mathcal{F} \subseteq A \subseteq P(S), A$ has finite intersection property\}. Partially order the elements of $\Omega$ by inclusion. If $C$ is a chain in $\Omega$, then $\bigcup C$ is an upper bound of $C$ in $\Omega$. By Zorn’s lemma, $\Omega$ has a maximal element $\mathcal{M}$. Being maximal, $\mathcal{M}$ has the following properties:

1. If $W_1,W_2,\ldots,W_n \in \mathcal{M}$, then $\mathcal{M} \cup \{W_1 \cap W_2 \cap \cdots \cap W_n\} \subseteq \mathcal{M}$ in $\Omega$ and so $W_1 \cap W_2 \cap \cdots \cap W_n \in \mathcal{M}$.
2. If $A \subseteq S$ and $A \cap W \neq \emptyset$ for all $W \in \mathcal{M}$, then $\mathcal{M} \cup \{A\} \not\subseteq \mathcal{M}$ in $\Omega$ and so $A \not\in \mathcal{M}$.
3. If $W_1 \in \mathcal{M}$ and $W_1 \subseteq A \subseteq S$, then $A \cap W \subseteq W_1 \cap W \neq \emptyset$ for all $W \in \mathcal{M}$ and so by (2), $A \not\in \mathcal{M}$.

For every $\alpha \in A$, let $\pi_\alpha$ be the projection of $S$ onto $S_\alpha$. Then $\{\pi_\alpha(X) : X \in \mathcal{M}\}$ is a collection of subsets of $S_\alpha$ having the finite intersection property so is $\{\pi_\alpha(X) : X \in \mathcal{M}\}$. By the fact that $\bigcap_{i=1}^n \pi_\alpha_i(X) \supseteq \bigcap_{i=1}^n \pi_\alpha_i(X)$. Since $S_\alpha$ is compact, there exists $x_\alpha \in \bigcap_{i=1}^n \pi_\alpha_i(X) : X \in \mathcal{M}$. Let $x \in S$ such that $\pi_\alpha(x) = x_\alpha$ for all $\alpha \in A$. To finish, we will show $x \in \overline{X}$ for every $X \in \mathcal{M}$. (Then $\mathcal{M}$ contains $\mathcal{F}$ and elements of $\mathcal{F}$ are closed simply $x \in \bigcap \mathcal{F}$. So $\bigcap \mathcal{F} \neq \emptyset$.)

For every open neighborhood $U$ of $x$ in $S$, there are open sets $U_\alpha$ in $S_\alpha$, $i = 1,2,\ldots,n$ such that $x \in \bigcap_{i=1}^n \pi_\alpha_i^{-1}(U_\alpha) \subseteq U$. Then $x_\alpha \in U_\alpha$. Since $x_\alpha \in \pi_\alpha_i(X)$ and $U_\alpha$ is an open neighborhood of $x_\alpha$, by the lemma, $U_\alpha$ intersects $\pi_\alpha(X)$ for every $X \in \mathcal{M}$. Then $\pi_\alpha^{-1}(U_\alpha)$ intersects every $X \in \mathcal{M}$. By (2), $\pi_\alpha^{-1}(U_\alpha) \in \mathcal{M}$. By (3), $\bigcap \mathcal{F} \neq \emptyset$. Then $U$ intersects every $X \in \mathcal{M}$. By the lemma, $x \in \overline{X}$ for every $X \in \mathcal{M}$. □

Remark. In 1950, John Kelley proved that Tychonoff’s theorem was equivalent to the axiom of choice.

Theorem (Banach-Alaoglu). Let $X$ be a normed space. The closed unit ball $B$ of $X^*$ is $w^*$-compact, i.e. $B$ is compact in the weak-star topology.
For each $x \in X$, let $D_x$ be the closed disk with center 0 and radius $\|x\|$ in $K$. By Tychonoff’s theorem, $D = \prod_{x \in X} D_x$ is compact. For $x \in X$ and $d \in D$, let $d_x$ denote the $x$-coordinate of $d$, i.e., $d_x = \pi_x(d)$. For every $y \in B$ and $x \in X$, since $\|y\| \leq 1$, $\|y(x)\| \leq \|y\| \|x\| \leq \|x\|$. So we may define $f : B \to D$ by letting $f(y) \in D$ to satisfy $f(y)_x = \pi_x(f(y)) = y(x)$ for all $x \in X$. Now $f$ is injective because $f(y_1) = f(y_2)$ implies for all $x \in X$, $y_1(x) = f(y_1)_x = f(y_2)_x = y_2(x)$, i.e., $y_1 = y_2$. Also, $f$ is a homeomorphism from $B$ (with the relative $w^*$-topology) onto $f(B)$ (with the relative product topology) because

$$\{z_\alpha\}_{\alpha \in I} \subset B \iff \forall x \in X, \{z_\alpha(x)\}_{\alpha \in I} \subset z(x) \in K \quad \text{(by property 3 of $w^*$-topology)}$$

$$\iff \forall x \in X, \{\pi_x(f(z_\alpha))\}_{\alpha \in I} \subset \pi_x(f(z)) \in K \quad \text{(by definition of $f(z_\alpha)$)}$$

$$\iff \{f(z_\alpha)\}_{\alpha \in I} \subset f(z) \in f(B) \quad \text{(by theorem on page 7).}$$

To see $B$ is $w^*$-compact, it is enough to show $f(B)$ is closed (hence compact) in $D$. Note

$$f(B) = \{w \in D : \forall a, b \in X, w_{a+b} = w_a + w_b \text{ and } \forall c \in K, x \in X, w_{cx} = cw_x\}.$$

Suppose $\{f(y_{\beta})\}_{\beta \in I} \to w \in D$. Then for every $x \in X$, since $\pi_x$ is continuous, we have $f(y_{\beta})_x \to w_x \in D_x$. So, for every $a, b, h, x \in X$ and $c \in K$, $f(y_{\beta})_{a+b} = f(y_{\beta})_a + f(y_{\beta})_b$ implies $w_{a+b} = w_a + w_b$ and $f(y_{\beta})_{cx} = cf(y_{\beta})_x$ implies $w_{cx} = cw_x$. Therefore, $w \in f(B)$.

**Remarks.** Using the Krein-Milman theorem and the Banach-Alaoglu theorem, it follows that the Banach spaces $C([0,1],\mathbb{R}), L^1([0,1])$, are not dual spaces of Banach spaces since their closed unit balls have too few extreme points and hence, the closed unit balls cannot be the closed convex hulls of the extreme points, see [Be], p. 110.

**Theorem (Helly).** Let $X$ be a Banach space. If $X$ is separable, then the closed unit ball $B$ of $X^*$ is $w^*$-sequentially compact (and hence all bounded sequences in $X^*$ have $w^*$-convergent subsequences.)

**Proof.** Let $S$ be a countable dense subset of $X$. Let $g_0 \in B$. By a diagonalization argument (as in the proof of the Arzela-Ascoli theorem), there is a subsequence $g_{n_k}$ such that $\lim_{k \to \infty} g_{n_k}(s)$ for all $s \in S$. Next, for every $x \in X$, we will show $\{g_{n_k}(x)\}$ is a Cauchy sequence, hence it converges. This is because for every $\varepsilon > 0$, there are $s \in S$ such that $\|x - s\| < \varepsilon/3$ and $N \in \mathbb{N}$ such that $j, k \geq N$ implies $\|g_{n_k}(s) - g_{n_j}(s)\| < \varepsilon/3$. So $j, k \geq N$ implies

$$\|g_{n_k}(x) - g_{n_j}(x)\| \leq \|g_{n_k}(x) - g_{n_k}(s)\| + \|g_{n_k}(s) - g_{n_j}(s)\| + \|g_{n_j}(s) - g_{n_j}(x)\|$$

$$\leq \|g_{n_k}\| \|x - s\| + \|g_{n_k}(s) - g_{n_j}(s)\| + \|g_{n_j}\| \|s - x\|$$

$$< 1(\varepsilon/3) + (\varepsilon/3) + 1(\varepsilon/3) = \varepsilon.$$

By part (a) of the Banach-Steinhaus theorem, $g(x) = \lim_{k \to \infty} g_{n_k}(x) \in X^*$ and $\|g\| \leq \liminf_{k \to \infty} \|g_{n_k}\| \leq 1$. By property (3) of weak-star topologies, we have $g_{n_k} \to^* g \in B$ in $X^*$.

**Remarks.** The converse of the theorem is false. If $X$ is a nonseparable Hilbert space, then by the Eberlein-Smulian theorem in the next section, the closed unit ball of $X^*$ is still $w^*$-sequentially compact.

§4. **Reflexivity.** Next we may inquire when the closed unit ball $B$ of a normed space $X$ is $w$-compact. To answer this, let $B^{**}$ be the closed unit balls of $X^{**}$. We have the following theorem.

**Theorem (Goldstine).** Let $X$ be a normed space. Then $B^{**} = \overline{i(B)}^{w^*}$, where $i$ is the canonical embedding. (Hence, $i(X)$ is $w^*$-dense in $X^{**}$ because $X^{**} = \bigcup_{n=1}^{\infty} nB^{**} = \bigcup_{n=1}^{\infty} \overline{i(nB)}^{w^*} = i(X)^{w^*}$.)

**Proof.** By the Banach-Alaoglu theorem, $B^{**}$ is $w^*$-closed, hence $B^{**} \supseteq \overline{i(B)}^{w^*}$. Assume there is $y \in B^{**} \setminus \overline{i(B)}^{w^*}$. Since $i(B)^{w^*}$ is convex and $w^*$-closed, by the separation theorem, there is a $w^*$-continuous
linear functional \( g \) on \( X^{**} \) such that \( \text{Reg}(y) < \inf \{ \text{Reg}(u) : u \in \iota(B)^{**} \} \). By the weak-star functional theorem, for all \( u \in X^{**} \), \( f(u) = -g(u) = u(z) \) for some \( z \in X^* \). Observe that there is \( c \in \mathbb{K} \) with \( |c| = 1 \) such that \( |z(x)| = z(cx) = \text{Re} z(cx) \). We have

\[
\|f\| \geq |f(y)| \geq \text{Re} f(y) > \sup \{ \text{Re} f(u) : u \in \iota(B)^{**} \} \geq \sup \{ \text{Re} u(z) : u = i_x \in \iota(B) \} = \sup \{ \text{Re} z(x) : x \in B \} \geq \sup \{ |z(x)| : x \in B \} = \|z\| = \|f\|.
\]

Then \( \|y\| > 1 \), i.e. \( y \notin B^{**} \), a contradiction. Therefore, \( B^{**} = \overline{i(B)}^{**} \).

**Remarks.**

1. We have \( i(B) = B^{**} \) if and only if \( i(X) = X^{**} \). This is because \( i(B) = B^{**} \) implies \( i(X) = \overline{i(B)} = B^{**} = X^{**} \) and conversely, if \( i(X) = X^{**} \), then for all \( f \in B^{**} \subseteq X^{**} = i(X) \), we have \( f = i_x \) for some \( x \in X \) (with \( \|x\| = \|f\| \leq 1 \) due to \( i \) is an isometry) so that \( f \in i(B) \).

2. The canonical embedding \( i : X \to i(X) \) is a homeomorphism when we take the \( w \)-topology on \( X \) and the \( w^* \)-topology on \( X^{**} \). This is because it is bijective and

\[
x_{\alpha} \xrightarrow{w} x \iff \forall f \in X^*, \ f(x_{\alpha}) \to f(x) \iff \forall f \in X^*, \ i_{x_{\alpha}}(f) \to i_x(f) \iff i_{x_{\alpha}} \xrightarrow{w^*} i_x.
\]

**Theorem (Banach-Smulian).** A normed space is reflexive iff its closed unit ball \( B \) is \( w \)-compact.

**Proof.** By the remarks and Goldstine’s theorem, \( B \) is \( w \)-compact in \( X \) iff \( i(B) \) is \( w^* \)-compact (hence \( w^* \)-closed) in \( X^{**} \) and \( B^{**} \) iff \( i(B) = \overline{i(B)}^{w^*} = B^{**} \) iff \( i(X) = X^{**} \).

Now reflexive spaces are dual spaces, hence they are complete. Which Banach spaces are reflexive? Also, observe that in addition to the \( w^* \)-topology on \( X^* \), there is also the weak topology on \( X^* \). Since \( \{ |f| : f \in X^{**} \} \supseteq \{ |i_x| : x \in X \} \), so the weak-star topology on \( X^* \) is a subset of the weak topology (which is a subset of the norm topology) on \( X^* \). Hence, on \( X^* \), \( w^* \)-open sets are \( w \)-open, \( w^* \)-closed sets are \( w \)-closed, but \( w \)-compact sets are \( w^* \)-compact. When are the \( w \)-topology and \( w^* \)-topology equal in \( X^{**} \)? The following theorem will answer both questions.

**Theorem.** Let \( X \) be a Banach space. The following are equivalent.

(a) \( X \) is reflexive.

(b) On \( X^* \), the weak topology is the same as the weak-star topology.

(c) \( X^* \) is reflexive.

**Proof.** (a) \( \Rightarrow \) (b) By (a), \( \{ |f| : f \in X^{**} \} = \{ |i_x| : x \in X \} \). So both topologies are generated by the same seminorms.

(b) \( \Rightarrow \) (c) By the Banach-Alaoglu theorem, the closed unit ball \( B^* \) of \( X^* \) is \( w^* \)-compact, hence \( w \)-compact by (b). By the Banach-Smulian theorem, \( X^* \) is reflexive.

(c) \( \Rightarrow \) (a) Since the canonical embedding is an isometry and the closed unit ball \( B \) of \( X \) is closed, hence complete, in \( X \), so \( i(B) \) is complete, hence closed, in \( X^{**} \). As \( i(B) \) is convex, by property (5) of weak topology, it is \( w \)-closed in \( X^{**} \). Since \( X^* \) is reflexive, applying (a) \( \Rightarrow \) (b) to \( X^* \), we see \( i(B) \) is also \( w^* \)-closed in \( X^{**} \). By Goldstine’s theorem, \( i(B) \overline{=} i(B)^{w^*} = B^{**} \). By remark (1) above, \( i(X) = X^{**} \).

**Theorem (Pettis).** If \( X \) is reflexive and \( M \) is a closed subspace of \( X \), then \( M \) is reflexive.

**Proof.** Let \( z \in M^{**} \). We have to show \( z = i_w \) for some \( w \in M \). Define \( T : X^* \to M^* \) by \( Tf = f|_M \). Since \( \|f|_M \| \leq \|f\| \), we get \( T \in L(X^*, M^*) \). Then \( z \circ T \in X^{**} = i(X) \). So there is \( w \in X \) such that \( z \circ T = i_w \), i.e. \( z(Tf) = f(w) \) for all \( f \in X^* \).

28
Assume \( w \in X \setminus M \). By the Hahn-Banach theorem, there is \( g \in X^* \) such that \( g(M) = 0 \) and \( g(w) = 1 \). Then \( Tg = g|_M = 0 \). However, \( 0 = z(Tg) = g(w) = 1 \), a contradiction. Hence \( w \in M \). Now for every \( h \in M^* \), by the Hahn-Banach theorem, there exists \( H \in X^* \) extending \( f \) (i.e. \( TH = H|_M = h \)). Then \( z(h) = z(TH) = H(w) = h(w) \) for all \( h \in M^* \). Therefore, \( z = i_w \).

Exercise. Prove that \( X \) is reflexive if and only if for any closed subspace \( M \) of \( X \), \( M \) and \( X/M \) are reflexive. See [KR], pp. 8-9.

Clearly every finite dimensional normed space is reflexive as its closed unit ball is compact, hence weak compact. If \( X \) is an infinite dimensional normed space, must \( X \) have some reflexive closed linear subspaces, other than the finite dimensional subspaces? The answer turns out to be negative. Below, we will show the only reflexive subspaces of \( \ell^1 \) are the finite dimensional subspaces. First, we need two theorems.

**Theorem (Banach).** For a normed space \( X \), if \( X^* \) is separable, then \( X \) is separable.

**Proof.** \( X = \{0\} \) is a trivial case. For \( X \neq \{0\} \), let \( D \) be a countable dense subset of \( X^* \). For every \( f \in D \), by the definition of \( \|f\| \) and the supremum property, there is \( x_f \in X \) such that \( \|x_f\| = 1 \) and \( |f(x_f)| \geq \|f\|/2 \). Let \( S \) be the set of all finite linear combinations of the \( x_f \)'s with rational coefficients. Then \( S \) is countable.

Next we will show \( S \) is dense in \( X \). By part (b) of the Hahn-Banach theorem, it suffices to show \( F \in X^* \) satisfying \( F \equiv 0 \) on \( S \) must be the zero functional. Since \( D \) is dense in \( X^* \), there exists a sequence \( \{f_n\} \) in \( D \) converging to \( F \). We have \( \|f_n - F\| \geq |(f_n - F)(x_f)| = |f_n(x_f)| \geq \|f_n\|/2 \), which implies \( \|f_n\| \to 0 \). Then \( F = 0 \).

**Remarks.** The converse is false in general. For example, \( \ell^1 \) is separable, but \( (\ell^1)^* = \ell^\infty \) is not separable. However, if \( X \) is a reflexive and separable Banach space, then since \( i \) is an isometry, \( X^{**} = i(X) \) is separable and hence \( X^* \) is separable by Banach’s theorem.

**Theorem (Eberlein-Smulian).** If \( X \) is reflexive, then the closed unit ball \( B \) of \( X \) is \( w\)-sequentially compact (and hence all bounded sequences in \( X \) have \( w\)-convergent subsequences).

**Proof.** Let \( \{x_n\} \) be a sequence in \( B \). Let \( M \) be the closed linear span of \( \{x_n\} \). By Pettis’ theorem, \( M \) is reflexive. Also \( M \) is separable as the set of all finite linear combinations of \( \{x_n\} \) with rational coefficients is dense. By the remark above, \( M^* \) is separable. By Helly’s theorem, \( \{ix_{n_k}\} \) in the closed unit ball \( B^{**} \) of \( M^{**} \) has a \( w^* \)-convergent subsequence \( \{ix_{n_{k_n}}\} \). By remark (2) before the Banach-Smulian theorem, \( \{x_{n_k}\} \) is a \( w \)-convergent subsequence of \( \{x_n\} \) in \( M \), say \( x_{n_k} \xrightarrow{w} x \in M \). For all \( f \in X^* \), we have \( f|M \in M^* \). By property 3 of weak topology, \( f(x_{n_k}) = f|M(x_{n_k}) \to f|M(x) = f(x) \), i.e. \( \{x_{n_k}\} \) \( w \)-converges to \( x \) in \( X \).

**Remarks.** In fact, Eberlein-Smulian proved a much deeper theorem, namely on any normed space (not necessarily reflexive), a subset is \( w \)-compact if it is \( w \)-sequentially compact. See [M], pp. 248-250.

Next, let \( M \) be a reflexive closed linear subspace of \( X = \ell^1 \). By the Eberlein-Smulian theorem, the closed unit ball of \( M \) is \( w \)-sequentially compact. Schur’s lemma below asserts that every \( w \)-convergent sequence in \( \ell^1 \) is convergent in the norm topology of \( \ell^1 \). Hence, the closed unit ball of \( M \) would be compact. By Riesz’ lemma, \( M \) would be finite dimensional. This implies \( \ell^1 \) is not reflexive and its only reflexive subspaces are the finite dimensional subspaces.

**Theorem (Schur’s Lemma).** If \( \{x^{(n)}\} \) is \( w \)-convergent in \( \ell^1 \), where \( x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)} \ldots) \) for \( n = 1, 2, 3, \ldots \), then \( \{x^{(n)}\} \) is convergent in the norm topology of \( \ell^1 \).

**Proof.** (Sliding Hump Argument) Assume \( x^{(n)} \xrightarrow{w} x \) in \( \ell^1 \), but \( x^{(n)} \to x \) is false. Replacing \( x^{(n)} \) by \( x^{(n)} - x \) if necessary, we may assume \( x = 0 \). Since \( \|x^{(n)}\|_1 \to 0 \) is false, passing to a subsequence, we may assume there is an \( \varepsilon > 0 \) such that \( \|x^{(n)}\|_1 = \sum_{j=1}^{\infty} |x_j^{(n)}| > \varepsilon \) for \( n = 1, 2, \ldots \).
Since \( x^{(n)} \to 0 \) in \( \ell^1 \), by property 3 of weak topology, \( \langle x^{(n)}, z \rangle = \sum_{j=1}^{\infty} z_j x_j^{(n)} \to 0 \) as \( n \to \infty \) for every \( z = (z_1, z_2, z_3, \ldots) \in \ell^\infty = (\ell^1)^* \). Our goal is to construct a special \( z \) with all \( |z_j| \leq 1 \) to get a contradiction of the last sentence.

First, by taking \( z = (0, \ldots, 0, 1, 0, \ldots) \), where 1 is in the \( j \)-th coordinate, we have for all \( j = 1, 2, 3, \ldots \), \( x_j^{(n)} \to 0 \) as \( n \to \infty \).

Next, define sequences \( \{m_k\}, \{n_k\} \) as follows. Set \( m_0 = 1, n_0 = 0 \). Inductively, for \( k \geq 1 \), suppose \( m_{k-1} \) and \( n_{k-1} \) are determined. By the last paragraph, \( \lim_{n \to \infty} \sum_{j=1}^{m_{k-1}} |x_j^{(n)}| = \sum_{j=1}^{m_{k-1}} \lim_{n \to \infty} |x_j^{(n)}| = 0 \). So we may let \( n_k \) be the smallest integer \( n > n_{k-1} \) such that \( \sum_{j=1}^{m_{k-1}} |x_j^{(n)}| < \frac{\varepsilon}{5} \). Since \( \sum_{j=1}^{\infty} |x_j^{(n_k)}| = \|x^{(n_k)}\|_1 < \infty \), we may let \( m_k \) be the smallest integer \( m > m_{k-1} \) such that \( \sum_{j=m_{k-1}+1}^{\infty} |x_j^{(n_k)}| < \frac{\varepsilon}{5} \).

Now observe that \( 1 = m_0 < m_1 < m_2 < \cdots \). Recall that sgn \( \alpha \) is the signum function defined to be \( |\alpha|/\alpha \) if \( \alpha \neq 0 \) and 1 if \( \alpha = 0 \). Let \( z = (z_1, z_2, \ldots) \in \ell^\infty \) be defined by \( z_1 = 0 \) and for \( k = 1, 2, 3, \ldots \), \( z_j = \text{sgn} \, x_j^{(n_k)} \), where \( m_{k-1} < j \leq m_k \). By the conditions on \( n_k \) and \( m_k \), we have \( z_j x_j^{(n_k)} = x_j^{(n_k)} \) for \( m_{k-1} < j \leq m_k \). So

\[
\sum_{j=1}^{\infty} |x_j^{(n_k)}| - z_j x_j^{(n_k)} \leq 2 \sum_{j=1}^{m_{k-1}} |x_j^{(n_k)}| + 2 \sum_{j=m_{k-1}+1}^{\infty} |x_j^{(n_k)}| < \frac{4\varepsilon}{5}.
\]

For \( k = 1, 2, 3, \ldots \), this gives

\[
\sum_{j=1}^{\infty} z_j x_j^{(n_k)} \geq \sum_{j=1}^{\infty} |x_j^{(n_k)}| - \sum_{j=1}^{\infty} (|x_j| - z_j x_j^{(n_k)}) > \varepsilon - \frac{4\varepsilon}{5} = \frac{\varepsilon}{5},
\]

which is a contradiction to \( \sum_{j=1}^{\infty} z_j x_j^{(n)} \to 0 \) as \( n \to \infty \).

Here is the reason why the proof is called a sliding hump argument. For each \( x^{(n_k)} \in \ell^1 \), if we plot the graph of \( f_{n_k}(x) = \sum_{j=1}^{\infty} |x_j^{(n_k)}| \chi_{(j-1,j)}(x) \) on the coordinate plane, then the area under the curve is greater than \( \varepsilon \) and the areas under the curve on \((0, m_{k-1}] \) and \((m_k, \infty) \) are both less than \( \varepsilon/5 \) so that the area under the curve on \((m_{k-1}, m_k) \) is greater than \( 3\varepsilon/5 \). Thus, we can say there is a hump in the middle portion over \((m_{k-1}, m_k) \). As \( k \) takes on the values \( 1, 2, 3, \ldots \), since \( 1 = m_0 < m_1 < m_2 < \cdots \), the humps of \( f_{n_k}(x) \) start to slide along the intervals \((m_0, m_1], (m_1, m_2], (m_2, m_3], \ldots \). Since the union of these intervals is \((0, \infty) \), we can patch up the sgn \( x_j^{(n_k)} \) on the intervals to get a \( z \in \ell^\infty \) to get a contradiction.
Chapter 4. Duality and Adjoints.

In this chapter, we introduce the adjoint operators. Also, we study how certain properties, such as surjectivity, density of ranges or closedness of ranges of operators can be expressed equivalently in terms of adjoint operators.

**Definitions.** For a nonempty subset $M$ of a normed space $X$, the annihilator of $M$ is

$$M^\perp = \{ y \in X^* : \langle x, y \rangle = 0 \text{ for all } x \in M \} = \bigcap_{x \in M} \ker i_x,$$

which is $w^*$-closed and norm-closed. For a nonempty subset $N$ of $X^*$, the (pre)annihilator of $N$ is

$$\perp N = \{ x \in X : \langle x, y \rangle = 0 \text{ for all } y \in N \} = \bigcap_{y \in N} \ker y,$$

which is $w$-closed and norm-closed.

**Remarks.** (1) For all $y \in X^*$, since $y \equiv 0$ on $M$ iff $y \equiv 0$ on $\overline{M}$, so $M^\perp = \overline{M}^\perp$. Similarly, $\perp N = \overline{N}^\perp$.

(2) By definitions above, $\{0\}^\perp = X^*$, $\perp \{0\} = X$, $X^\perp = \{0\}$. Also, $\perp (X^*) = \{0\}$, where the left-to-right inclusion uses part (c) of the Hahn-Banach theorem.

**Notations.** For a subset $M$ of $X$, we write $M^+\perp$ to mean $+M^\perp$. For a subset $N$ of $X^*$, we write $N^\perp\perp$ to mean $(\perp N)^\perp$. From definitions above, we have $M \subseteq M^{+\perp} \subseteq X$ and $N \subseteq N^{\perp\perp} \subseteq X^*$.

Although in the definitions of annihilator and preannihilator, $M$ and $N$ may be any nonempty subset of the normed space, in the sequel, we will only consider the cases $M$ and $N$ are vector subspaces.

**Double-Perp Theorem.** Let $X$ be a normed space.

(a) If $M$ is a vector subspace of $X$, then $M^{+\perp} = \overline{M^\perp}$, the norm-closure or weak-closure of $M$.

(b) If $N$ is a vector subspace of $X^*$, then $N^{\perp\perp} = \overline{N^{w^*}}$, the weak-star closure of $N$.

**Proof.** (a) Since $M \subseteq M^{+\perp}$, so $\overline{M} \subseteq M^{+\perp}$, Assume there is $x \in M^{+\perp} \setminus \overline{M}$. By part (b) of the Hahn-Banach theorem, there is $y \in X^*$ such that $y \equiv 0$ on $M$ and $y(x) \neq 0$. So $y \in M^\perp$ and $x \notin M^{+\perp}$, a contradiction. Therefore, $\overline{M} = M^{+\perp}$.

(b) Since $N \subseteq N^{\perp\perp}$, so $\overline{N^{w^*}} \subseteq N^{\perp\perp}$. Assume there is $y \in N^{\perp\perp} \setminus \overline{N^{w^*}}$. Applying part (b) of the corollary to the separation theorem to $X^*$ with the $w^*$-topology and the weak-star functional theorem, there is $w^*$-continuous linear functional $g = i_x$ on $X^*$ such that $g = i_x \equiv 0$ on $N$ and $y(x) = g(y) \neq 0$. So $x \notin \perp N$ and $y \notin N^{\perp\perp}$, a contradiction. Therefore, $\overline{N^{w^*}} = N^{\perp\perp}$. \[\square\]

**Remarks.** We have $M^\perp = \{0\}$ iff $\overline{M} = X$, which can be checked by taking (pre)annihilators of both sides. Similarly, $M^\perp = X^*$ iff $M = \{0\}$; \ $\perp \overline{N^{w^*}} = X^*$; \ $\perp \overline{N} = X$ iff $N = \{0\}$.

**Duality Theorem.** Let $M$ be a closed vector subspace of a normed space $X$. We have the following isometric isomorphisms and equations.

(a) $M^* \cong X^*/M^\perp$. For every $F \in X^*$, $\sup \{|\langle x, F \rangle | : x \in M, \|x\| \leq 1\} = \min \{\|F - G\| : G \in M^\perp\}$.

(b) $(X/M)^* \cong M^\perp$. For every $x \in X$, $\inf \{|\langle x - m \rangle | : m \in M\} = \max \{|\langle x, G \rangle | : G \in M^\perp, \|G\| \leq 1\}$.

**Proof.** (a) Define $\phi : M^* \to X^*/M^\perp$ by $\phi(f) = F + M^\perp$, where $F \in X^*$ is any linear extension of $f \in M^*$. (If $F$ and $F'$ are linear extensions of $f$, then $F - F' \in M^\perp$ so that $F + M^\perp = F' + M^\perp$. Hence $\phi$ is well defined.) Clearly $\phi$ is linear. For every $F \in X^*$, we have $\phi(F|_M) = F + M^\perp$. So $\phi$ is surjective.
Now we show $\phi$ is isometric. For $f \in M^*$, by part (a) of the Hahn-Banach theorem, there is a linear extension $F' \in X^*$ of $f \in M^*$ such that $\|F'\| = \|f\|$. For every $G \in M^\perp$, $F' - G$ also linearly extends $f$, so $\|f\| \leq \|F' - G\|$. Then $\|f\| \leq \|F' + M^\perp\| = \inf\{\|F' - G\| : G \in M^\perp\} \leq \|F'\| = \|f\|$. (Thus, there is equality throughout and so the infimum attained by $F'$.) We have $\|f\| = \|F' + M^\perp\| = \|\phi(f)\|$

Finally, for every $F \in X^*$, taking $f = F|_M$, we have $F = F'$ on $M$. So $\|F|_M\| = \|f\| = \|F' + M^\perp\| = \|F + M^\perp\|$, which is the equation in the second sentence.

(b) Recall the quotient mapping $\pi : X \to X/M$ is defined by $\pi(x) = x + M$. Define $\tau : (X/M)^* \to M^\perp$ by $\tau(F) = F \circ \pi$ and call this $f$. (Now $\pi \in L(X, X/M)$ and $F \in (X/M)^*$ imply $f \in X^*$. If $x \in M$, then $f(x) = F(x + M) = F([0]) = 0$ and so $f \in M^\perp$.) Clearly, $\tau$ is linear.

Next we will show $\tau$ is surjective and isometric. For every $f \in M^\perp$, since $M \subseteq \ker f$, the function $F : X/M \to \mathbb{K}$ given by $F(x + M) = f(x)$ is well-defined and linear. The key step is to show $\|F\| = \|f\|$. This gives $F \in (X/M)^*$, $\tau(F)$ is linear and $\tau$ is an isometric isomorphism.

For all $m \in M$, $|F(x + M)| = |f(x)| = |f(x - m)| \leq \|f\| \|x - m\|$. Taking infimum over all $m \in M$, $|F(x + M)| \leq \|f\| \|x + M\|$. Then $\|F\| \leq \|f\|$ (and so $F$ is continuous). Also, $|f(x)| = |F(x + M)| \leq \|F\| \|x + M\| \leq \|F\| \|x\|$. So $\|F\| = \|f\| = \|\tau(F)\|$

For the equation in the second sentence, let $x \in X$. By part (c) of the Hahn-Banach theorem, there is $F_0 \in (X/M)^*$ such that $\|F_0\| = 1$ and $F_0(x + M) = \|x + M\|$. Let $f_0 = \tau(F_0) \in M^\perp$, then $f_0(x) = \tau(F_0)(x) = F_0(x + M) = \|x + M\|$. Also, $\tau$ isometric implies $\|f_0\| = \|F_0\| = 1$.

For all $G \in M^\perp$, $\|G\| \leq 1$ and $m \in M$, we have $\|\langle x, G \rangle\| = \|G(x)\| = \|G(x - m)\| \leq \|x - m\|$. Since $f_0$ is such a $G$, we have $f_0(x) = \sup\{|\langle x, G \rangle| : G \in M^\perp, \|G\| \leq 1\} \leq \inf\{|\langle x - m, m \rangle| : m \in M\} = \|x + M\| = F_0(x + M) = f_0(x)$.

(Thus, there is equality throughout and the supremum is attained by $f_0(x) = |\langle x, f_0 \rangle|$.)

**Remarks.** If $M$ is a finite dimensional subspace of $X$, then $\dim M = \dim M^* = \dim(X/M^\perp) = \text{codim} M^\perp$ by (a). If $M$ is a closed subspace of finite codimension in $X$, then $\dim M = \dim(X/M) = \dim(X/M)^* = \dim M^\perp$.

**Definition.** Let $X, Y$ be normed spaces over $\mathbb{K}$. For $T \in L(X, Y)$ and $y \in Y^*$, define $T^* : Y^* \to X^*$ by $T^*(y) = y(T(x)) = y \circ T \in X^*$. Thus, for all $x \in X$, $\langle x, T^*(y) \rangle = y(T(x)) = \langle T(x), y \rangle$. $T^*$ is called the adjoint of $T$.

**Notations.** For convenience, we will write $T(x)$ as $Tx$ and $S \circ T$ as $ST$ when no confusion arises.

**Theorem (Properties of Adjoint Operators).** If $X, Y, Z$ are normed spaces over $\mathbb{K}$, $c_1, c_2 \in \mathbb{K}$, $S \in L(Y, Z)$ and $T_1, T_2 \in L(X, Y)$, then

(a) $\|T^*\| = \|T\|$ and hence $T^* \in L(Y^*, X^*)$

(b) $(c_1 T_1 + c_2 T_2)^* = c_1 T_1^* + c_2 T_2^*$

(c) $(S \circ T)^* = T^* \circ S^*$ and for the identity operator $I \in L(X), I^* = I$

(d) $T^{**} \in L(X^{**}, Y^{**})$ and identifying $X$ with $\text{im}(X) \subseteq X^{**}$, we have $T^{**}|_X = T$

(e) if $T$ is invertible, then $T^*$ is also invertible and $(T^*)^{-1} = (T^{-1})^* \in L(X^*, Y^*)$

(f) if $T^*$ is invertible, then $T$ is bounded below, hence injective. In case $X$ is a Banach space, $T^*$ invertible implies $T$ invertible and $Y$ complete.

**Proof.** (a) $\|T\| = \sup\{|\langle T(x), y \rangle| : \|x\| \leq 1\} = \sup\{|\langle T(x), y \rangle| : \|x\| \leq 1, \|y\| \leq 1\}$

(b) $(c_1 T_1 + c_2 T_2)^*(y) = y \circ (c_1 T_1 + c_2 T_2) = c_1 y \circ T_1 + c_2 y \circ T_2 = (c_1 T_1^* + c_2 T_2^*)(y)$.
(c) \((S \circ T)^* (y) = y \circ (S \circ T) = T^* (y \circ S) = T^* \circ S^* (y)\). \(I^* (y) (x) = y (I(x)) = y (x)\) for all \(x \in X\). So \(I^* (y) = y\).

(d) For \(x \in X\), \(T^* (x) = T^* (i_x) = i_x \circ T^* = i_{T(x)} = T(x)\).

(e) Applying (c) to \(T \circ T^{-1} = I\) and \(T^{-1} \circ T = I\), we get \((T^{-1})^* \circ T^* = I^* = I\) and \(T^* \circ (T^{-1})^* = I^* = I\).

So \((T^*)^{-1} = (T^{-1})^*\) and it is in \(L(X^*, Y^*)\) by the inverse mapping theorem.

(f) By (e), \(T^*\) invertible implies \(T^{**}\) invertible. Hence \(T^{**}\) is bounded below. By (d), \(T\) is bounded below, so \(T\) is injective.

In case \(X\) is a Banach space, by the lower bound theorem, \(T(X)\) is complete and hence closed. Assume there is \(y \in Y \setminus T(X)\). Then \(Y \neq \{0\}\). By part (b) of the Hahn-Banach theorem, there is \(F \in Y^*\) with \(\|F\| = 1\) such that for all \(x \in X\), \(0 = F(T(x)) = T^* (F(x))\), i.e. \(T^* (F) = 0\). Since \(T^*\) is invertible (in particular, injective), we get \(F = 0\), a contradiction. Hence \(T(\langle x, y \rangle) = Y\). Then \(Y = \text{ran } T\). Then \(T\) is complete and \(T\) is injective. By the inverse mapping theorem, \(T^{-1} \in L(Y, X)\) and \(T\) is invertible.

**Theorem (Kernel-Range Relations).** Let \(X, Y\) be normed spaces and \(T \in L(X, Y)\). Then

\[
\ker T = \perp (\text{ran } T^*), \quad \ker T^* = (\text{ran } T)^\perp, \quad (\ker T)^\perp = \text{ran } T^{w^*}\quad \text{and} \quad \perp (\ker T^*) = \text{ran } T.
\]

**Proof.** For the first equation,

\[
x \in \ker T \iff T(x) = 0 \iff \{T(x)\}^\perp = Y^* \iff \text{all } y \in Y^*, \quad 0 = \langle T(x), y \rangle = \langle x, T^* (y) \rangle \iff x \in \perp (\text{ran } T^*).
\]

For the second equation,

\[
y \in \ker T^* \iff T^* (y) = 0 \iff \perp (T^* (y)) = X \iff \text{all } x \in X, \quad 0 = \langle x, T^* (y) \rangle = \langle T(x), y \rangle \iff y \in (\text{ran } T)^\perp.
\]

For the third equation, by the first equation, \(\ker T = \perp (\text{ran } T^*)\). By the double-perp theorem, \((\ker T)^\perp = (\text{ran } T^*)^\perp = \text{ran } T^{w^*}\).

For the fourth equation, by the second equation, \(\ker T^* = (\text{ran } T)^\perp\). By the double-perp theorem, \(\perp (\ker T^*) = (\text{ran } T)^\perp = \text{ran } T\).

**Corollary 1.** Let \(X, Y\) be normed spaces and \(T \in L(X, Y)\). Then

(a) \(\ker T = (\ker T)^\perp\) and \(\ker T^* = (\ker T^*)^\perp\),

(b) \(\text{ran } T\) is dense (or \(w\)-dense) in \(Y\) iff \(T^*\) is injective,

(c) \(\text{ran } T^*\) is \(w^\perp\)-dense in \(X^*\) iff \(T\) is injective.

**Proof.** (a) \(\ker T = \perp (\text{ran } T^*)\) is norm-closed in \(X\) so that \(\ker T = \overline{\ker T} = (\ker T)^\perp\) and \(\ker T^* = (\text{ran } T)^\perp\) is \(w^\perp\)-closed in \(Y^*\) so that \(\ker T^* = \overline{\ker T^*} = (\ker T^*)^\perp\).

(b) If \(T^*\) is injective, then by the last two theorems, \(\overline{\text{ran } T} = \perp (\ker T^*) = \perp \{0\} = Y\). Conversely, if \(\text{ran } T\) is dense (equivalently, \(w\)-dense) in \(Y\), then by (a), \(\ker T^* = (\ker T^*)^\perp = \overline{\text{ran } T}^\perp = \overline{\{0\}} = \{0\}\).

(c) If \(T\) is injective, then by the last two theorems, \(\overline{\text{ran } T^{w^*}} = (\ker T)^\perp = \overline{\{0\}} = X^*\). Conversely, if \(\text{ran } T^*\) is \(w^\perp\)-dense in \(X^*\), then by (a), \(\ker T = (\ker T)^\perp = \perp (\overline{\text{ran } T^{w^*}}) = \perp (X^*) = \{0\}\).

**Corollary 2.** Let \(X\) be a Banach space, \(Y\) a normed space and \(T \in L(X, Y)\). The following are equivalent:

(a) \(T\) invertible,

(b) \(T^*\) invertible,

(c) \(T\) is bounded below and \(\text{ran } T\) is dense in \(Y\),

(d) \(T\) and \(T^*\) are both bounded below.
By properties (e) and (f) of the adjoint operators, we have (a) \( \iff \) (b). Next, (a) \( \iff \) (c) and also (a), (b) \( \Rightarrow \) (d) follow from the lower bound theorem. Finally, (d) \( \Rightarrow \) (c) due to (b) of the last corollary.

Closed Range Theorem. Let \( X, Y \) be Banach spaces and \( T \in L(X,Y) \). The following are equivalent.

(a) \( \text{ran } T \) is norm-closed (or \( w \)-closed),
(b) \( \text{ran } T^* \) is \( w^* \)-closed,
(c) \( \text{ran } T^* \) is norm-closed.

Proof. Let \( X_0 = X/\ker T \) and \( Y_0 = \overline{\text{ran } T} \). The map \( T_0 : X_0 \to Y_0 \) given by \( T_0(x + \ker T) = T(x) \) is well-defined, linear and injective. Also, \( \text{ran } T_0 = \text{ran } T \). Next we will compute \( T_0^* \). By the duality theorem, we have \( X_0^* = (\ker T)^\perp \) and \( Y_0^* = Y^*/(\text{ran } T)^\perp = Y^*/\ker T^* \). Now \( T_0^* : Y_0^* = Y^*/\ker T^* \to X_0^* = (\ker T)^\perp \) is given by \( T_0^*(\psi + \ker T^*) = T^*(\psi) \), which is well-defined, linear and injective. Also, \( \text{ran } T_0^* = \text{ran } T^* \).

(a) \( \Rightarrow \) (b) Since \( \text{ran } T \) is norm-closed, \( \text{ran } T_0 = \text{ran } T = \overline{\text{ran } T} = Y_0 \), i.e. \( T_0 \) is surjective (hence bijective). By the inverse mapping theorem, \( T_0^* \) is invertible. So \( T_0^* \) is invertible, hence surjective. So \( \text{ran } T^* = \text{ran } T_0^* = X_0^* = (\ker T)^\perp \) is \( w^* \)-closed.

(b) \( \Rightarrow \) (c) The weak-star topology is a subset of the norm topology.

(c) \( \Rightarrow \) (a) Since \( \text{ran } T^* \) is norm-closed, \( \text{ran } T_0^* \) is closed. Since \( T_0^* \) is also injective, by the lower bound theorem, \( T_0^* \) is bounded below. Hence there is \( \delta > 0 \) such that \( \| T_0^* (u) \| \geq \delta \| u \| \) for all \( u \in Y_0 \).

To show \( \text{ran } T \) is norm-closed, it suffices to show \( T_0 \) is open (as it would implies \( T_0 \) is surjective and hence \( \text{ran } T = \text{ran } T_0 = Y_0 = \overline{\text{ran } T} \)). Now to show \( T_0 \) is open, let \( U \) be the unit ball in \( X_0 \). It is enough to show \( T_0(U) \) is a neighborhood of \( 0 \) in \( Y_0 \). Using the lemma prior to the open mapping theorem, it is further enough to show \( T_0(U) \) contains the open ball \( B(0, \delta) \) with center at \( 0 \) and radius \( \delta \).

Let \( v \in Y_0 \setminus T_0(U) \). By the separation theorem, there is a \( g \in Y_0^* \) such that \( \text{Re } g(v) < \inf \{ \text{Re } g(T_0(u)) : u \in U \} \). Let \( f = -g/\| T_0^* g \| \), then \( \| T_0^* f \| = 1 \) and \( | T_0^* f(u) | = e^{i\theta} T_0^* f(u) = T_0^* f(e^{i\theta} u) = \text{Re } T_0^* f(e^{i\theta} u) \). So

\[
\text{Re } f(v) > \sup \{ \text{Re } f(T_0(u)) : u \in U \} = \sup \{ | T_0^* f(u) | : u \in U \} = \| T_0^* f \| = 1.
\]

As \( T_0^* \) is bounded below, we get \( 1 = \| T_0^* f \| \geq \delta \| f \| \). So \( \| f \| \leq \frac{1}{\delta} \). Now \( \frac{1}{\delta} \| v \| \geq \| f \| \| v \| \geq | f(v) | \geq \text{Re } f(v) > 1 \). We get \( \| v \| > \delta \). Hence, \( v \notin B(0, \delta) \). Therefore, \( T_0(U) \) contains \( B(0, \delta) \).

Corollary. Let \( X, Y \) be Banach spaces and \( T \in L(X,Y) \). Then \( T \) is surjective iff \( T^* \) is bounded below. Similarly, \( T^* \) is surjective iff \( T \) is bounded below.

Proof. \( T \) is surjective iff \( \text{ran } T \) is dense and norm-closed in \( Y \). By corollary 1 and the closed range theorem, this is iff \( T^* \) is injective and \( \text{ran } T^* \) is norm-closed. By the lower bound theorem, this is iff \( T^* \) is bounded below. The second statement can be proved similarly.
Chapter 5. Basic Operator Facts on Banach Spaces.

§1. Spectrum. We will study operators in Banach spaces over \( \mathbb{C} \) in this chapter. So all vector spaces refered to below when not specified will mean Banach spaces over \( \mathbb{C} \). We begin with the observation that for a Banach space \( X \), \( L(X) = L(X, X) \) is not only a Banach space, but it has a continuous multiplication structure.

**Definition.** A Banach algebra is a Banach space with a multiplication such that \( \|xy\| \leq \|x\| \cdot \|y\| \) for all \( x \) and \( y \) in the space. (Note \( x_n \to x \) and \( y_n \to y \) implies \( \|x_n\|, \|y_n\| \) bounded and
\[
\|x_n y_n - xy\| = \|x_n(y_n - y) + (x_n - x)y\| \leq \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\| \to 0.
\]
So multiplication is continuous.)

**Example.** Let \( X, Y, Z \) be normed spaces, \( T \in L(X, Y) \) and \( S \in L(Y, Z) \), then \( S \circ T \in L(X, Z) \). For every \( x \in X \), \( \|(S \circ T)(x)\| = \|S(T(x))\| \leq \|S\| \|T(x)\| \leq \|S\| \|T\| \|x\| \) Thus, \( \|S \circ T\| \leq \|S\| \|T\| \).

In the case \( X = Y = Z \) is a Banach space, this asserts that \( L(X) \) is a Banach algebra with composition as multiplication.

As in linear algebra, for an operator \( T \in L(X) \), the related operator \( T - cI \) is important.

**Definitions.** Let \( X \) be a Banach space over \( \mathbb{C} \) and \( T \in L(X) \).

1. The **resolvent set** of \( T \) is \( \rho(T) = \{c \in \mathbb{C} : T - cI \text{ is invertible}\} \). For \( c \in \rho(T) \), the operator \( R_c(T) = (cI - T)^{-1} \) is called the **resolvent** of \( T \).
2. The **spectrum** of \( T \) is \( \sigma(T) = \{c \in \mathbb{C} : T - cI \text{ is non-invertible}\} \). A common alternative notation is \( sp(T) \).
3. The **point spectrum** of \( T \) is the set \( \sigma_p(T) = \{c \in \mathbb{C} : \ker(T - cI) \neq \{0\}\} \) of eigenvalues of \( T \).
4. The **approximate point spectrum** is the set \( \sigma_{ap}(T) = \{c \in \mathbb{C} : T - cI \text{ is not bounded below}\} = \{c \in \mathbb{C} : \exists x_1, x_2, x_3, \ldots \in X, \|x_i\| = 1, (T - cI)(x_i) \to 0\} \) of all approximate eigenvalues of \( T \).
5. The **compression spectrum** of \( T \) is the set \( \sigma_{com} = \{c \in \mathbb{C} : \text{ran}(T - cI) \neq X\} \).
6. The **residual spectrum** of \( T \) is the set \( \sigma_r(T) = \sigma_{com}(T) \setminus \sigma_p(T) = \{c \in \mathbb{C} : \ker(T - cI) = \{0\}, \text{ran}(T - cI) \neq X\} \).
7. The **continuous spectrum** of \( T \) is the set \( \sigma_c(T) = \sigma(T) \setminus (\sigma_p(T) \cup \sigma_{com}(T)) = \{c \in \mathbb{C} : \ker(T - cI) = \{0\}, \text{ran}(T - cI) \subset \text{ran}(T - cI) = X\} \).

**Remarks.** Since an operator is invertible iff it is injective and surjective (i.e. its range is closed and dense) iff it is bounded below and its range is dense, so \( \sigma(T) = \sigma_{ap}(T) \cup \sigma_{com}(T) \). Clearly, \( \sigma_p(T) \subseteq \sigma_{ap}(T) \), but \( \sigma_p(T) \cap \sigma_{com}(T) \) may not be empty (eg. \( T \) has rank 1) so that \( \sigma_{ap}(T), \sigma_{com}(T) \) may not be disjoint. To get disjoint decomposition of \( \sigma(T) \), we can write \( \sigma(T) \) as the union of the pairwise disjoint sets \( \sigma_p(T), \sigma_r(T), \sigma_c(T) \).

**Theorem.** For every operator \( T \in L(X) \), \( \sigma(T) = \sigma(T^*) \).

**Proof.** This follows easily from the fact that \( T - cI \) is invertible if and only if \( T^* - cI = (T - cI)^* \) is invertible on a Banach space \( X \).

**Lemma on Inverses.** (1) If \( T \in L(X) \) is invertible and \( S \in L(X) \) such that \( \|S\| < \|T^{-1}\|^{-1} \), then \( T - S \) is invertible. So the set of invertible operators in \( L(X) \) is an open set.

(2) The map \( T \mapsto T^{-1} \) on the set of invertible operators is continuous.

**Proof.** (1) Let \( R = T^{-1}S \), then \( \|R\| \leq \|T^{-1}\| \|S\| < 1 \) and \( \sum_{i=0}^{\infty} R^i \) converges absolutely in \( L(X) \). The sum is easily checked to be \((I - R)^{-1}\). Then \( T - S = T(I - R) \) is invertible.
(2) For $T$ invertible and $\|S\| < \|T^{-1}\|^{-1}$, let $R = T^{-1}S$. As $\|S\| \to 0$, $\|R\| \leq \|T^{-1}\|\|S\| \to 0$, which implies $\|(T - S)^{-1} - T^{-1}\| = \|(I - R)^{-1} - I\| \leq \sum_{i=1}^{\infty} R^i \|T^{-1}\| \leq \frac{|R|}{1 - \|R\|} \|T^{-1}\| \to 0$.

**Resolvent Identity.** $R_a(T) - R_b(T) = (b-a)R_a(T)R_b(T)$.

**Proof.** Let $A = aI - T = R_a(T)^{-1}$ and $B = bI - T = R_b(T)^{-1}$, then $B - A = (b-a)I$ and $A^{-1} - B^{-1} = A^{-1}BB^{-1} - A^{-1}AB^{-1} = A^{-1}(B - A)B^{-1} = (b-a)A^{-1}B^{-1}$.

**Remarks.** Two operators $T_0$ and $T_1$ are said to commute iff $T_0T_1 = T_1T_0$. The resolvent identity implies $R_a(T)$ and $R_b(T)$ commute since $R_a(T)R_b(T) = \frac{R_a(T) - R_b(T)}{b-a} = R_b(T)R_a(T)$. Also, $\lim_{a \to b} \frac{R_a(T) - R_b(T)}{a-b} = -R_b(T)^2$, the limit being taken in the norm of $L(X)$.

Concerning the spectrum of an operator, we have the following important facts.

**Gelfand’s Theorem.** For every $T \in L(X)$, $\sigma(T)$ is a nonempty compact set in $\mathbb{C}$.

**Gelfand-Mazur Theorem.** Let $r(T) = \max\{|z| : z \in \sigma(T)\}$. Then,

$$r(T) = \inf\{\|T^m\|^{1/m} : m = 1, 2, 3, \ldots\} = \lim_{m \to \infty} \|T^m\|^{1/m}.$$

($r(T)$ is the farthest distance of any point in $\sigma(T)$ from the origin and is called the spectral radius of $T$.)

Using these theorems, we will look at some examples first.

**Examples.** (1) Define the (backward) shift operator $T : \ell^1 \to \ell^1$ by $T(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots)$. Now,

$$\|T(x_1, x_2, x_3, \ldots)\|_1 = \|(x_2, x_3, x_4, \ldots)\|_1 = \sum_{i=2}^{\infty} |x_i| \leq \sum_{i=1}^{\infty} |x_i| = \|(x_1, x_2, x_3, \ldots)\|_1.$$

So $T$ is bounded, hence continuous. If $x_1 = 0$, then the above inequality becomes an equality. So $\|T\| = 1$. Since $r(T) = \lim_{n \to \infty} \|T^m\|^{1/n} \leq \|T\| = 1$, so $\sigma(T)$ is a nonempty compact subset of $B(0, 1) = \{z \in \mathbb{C} : |z| \leq 1\}$.

If $|z| < 1$, then $T(1, z, z^2, \ldots) = (z, z^2, z^3, \ldots) = z(1, z, z^2, \ldots)$. So $T - zI$ is not invertible as $(1, z, z^2, \ldots) \in \ker(T - zI)$. Then $B(0, 1) = \{z \in \mathbb{C} : |z| < 1\}$ is a subset of $\sigma(T)$. As $\sigma(T)$ is closed, $\sigma(T) = B(0, 1)$.

Define (forward or unilateral) shift operator $S : \ell^\infty \to \ell^\infty$ by $S(y_1, y_2, y_3, \ldots) = (0, y_1, y_2, y_3, \ldots)$. $S = T^*$ because from $(\ell^1)^* = \ell^\infty$ under the pairing $\langle (a_1, a_2, a_3, \ldots), (b_1, b_2, b_3, \ldots) \rangle = a_1b_1 + a_2b_2 + a_3b_3 + \cdots$, we have for all $(x_1, x_2, x_3, \ldots) \in \ell^1$,

$$\langle (x_1, x_2, x_3, \ldots), T^*(y_1, y_2, y_3, \ldots) \rangle = \langle T(x_1, x_2, x_3, \ldots), (y_1, y_2, y_3, \ldots) \rangle = x_2y_1 + x_3y_2 + x_4y_3 + \cdots$$

$$= \langle (x_1, x_2, x_3, \ldots), (0, y_1, y_2, y_3, \ldots) \rangle = \langle (x_1, x_2, x_3, \ldots), S(y_1, y_2, y_3, \ldots) \rangle.$$

Now $\|S\| = \|T^*\| = \|T\| = 1$ and $\sigma(S) = \sigma(T^*) = \sigma(T) = B(0, 1)$.

(2) Define the Volterra operator $V : C[0, 1] \to C[0, 1]$ by $(Vf)(x) = \int_0^x f(t) \, dt$. We have

$$\|V\|_\infty = \sup_{x \in [0, 1]} \int_0^x |f(t)| \, dt \leq \int_0^1 |f(t)| \, dt \leq \|f\|_\infty.$$
So $V$ is bounded. For $f \equiv 1$, $(Vf)(x) = x$, $\|Vf\|_\infty = 1 = \|f\|_\infty$ and so $\|V\|_\infty = 1$. Next we show $|(V^n f)(x)| \leq \|f\|_\infty \frac{x^n}{n!}$ for all $x \in [0, 1]$. For $n = 1$, $|(Vf)(x)| = \left| \int_0^x f(t) \, dt \right| \leq \|f\|_\infty x$. Assuming case $n$, we have

$$|(V^{n+1} f)(x)| = \left| \int_0^x (V^n f)(t) \, dt \right| \leq \int_0^x \|f\|_\infty \frac{t^n}{n!} \, dt \leq \|f\|_\infty \frac{x^{n+1}}{(n+1)!}.$$ 

This implies $\|V^n f\|_\infty \leq \|f\|_\infty \frac{1}{n!}$. For $f \equiv 1$, we get equality. Hence $\|V^n\| = \frac{1}{n!}$. Since $\lim_{n \to \infty} \frac{1/(n+1)!}{1/n!} = \lim_{n \to \infty} \frac{1}{n+1} = 0$, we get $\lim_{n \to \infty} \left(\frac{1}{n!}\right)^{1/n} = 0$. So $r(V) = 0$ and $\sigma(V) = \{0\}$, but $\ker V = \{0\}$ implies $\sigma_p(V) = \emptyset$.

**Remarks.** If $\sigma(T) = \{0\}$, then $T$ is called a *quasinilpotent* operators. We can also define $V : L^2[0, 1] \to L^2[0, 1]$ by $(Vf)(x) = \int_{[0, x]} f \, dm$. Then $\|V\| = \frac{2}{\pi}$ and $\sigma(V) = \{0\}$. See [H], problems 186 to 188.

(3) For $f \in L^\infty[0, 1]$, define the *multiplication operator* $M_f : L^1[0, 1] \to L^1[0, 1]$ by $M_f(g) = fg$. We will show $\|M_f\| = \|f\|_\infty$. The case $f = 0$ is clear. So we consider $f \neq 0$ in $L^\infty[0, 1]$.

Clearly, $\|M_f(g)\|_1 = \left| \int_{[0, 1]} fg \, dm \right| \leq \int_{[0, 1]} \|fg\| \, dm \leq \|f\|_\infty \|g\|_1$. So $\|M_f\| \leq \|f\|_\infty$. Conversely, we may think of $f$ as a bounded measurable function on $[0, 1]$ (by taking a representative in the equivalence class of $f \in L^\infty[0, 1]$). Let $A_n = \{x \in [0, 1] : |f(x)| > \|f\|_\infty - \frac{1}{n}\}$ and $g_n = \frac{\chi_{A_n}}{\int m(A_n)}$. Then $\|g_n\|_1 = 1$ and

$$\|f\|_\infty - \frac{1}{n} \leq \frac{1}{m(A_n)} \int_{A_n} |f| \, dm = \int_{[0, 1]} |fg_n| \, dm \leq \|f\|_\infty \|g_n\|_1 = \|f\|_\infty.$$ 

So $\|M_f(g_n)\|_1 = \int_{[0, 1]} |fg_n| \, dm \to \|f\|_\infty$ as $n \to \infty$. Therefore, $\|M_f\| = \|f\|_\infty$.

For $\sigma(M_f)$, consider the *essential range* of $f$, which is $S = \{z \in \mathbb{C} : m(f^{-1}(B(z, r))) > 0$ for all $r > 0\}$. If $z \in S$, then let $D_n = f^{-1}(B(z, \frac{1}{n}))$ and $h_n = \frac{\chi_{D_n}}{m(D_n)}$. Then $\|h_n\|_1 = 1$ and

$$\|(M_f - zI)h_n\|_1 = \int_{[0, 1]} |f - z| h_n \, dm = \frac{1}{m(D_n)} \int_{D_n} |f - z| \, dm \leq \frac{1}{n}.$$ 

Assume $M_f - zI$ has an inverse $L$, then $1 = \|h_n\|_1 = \|L(M_f - zI)h_n\|_1 \leq \|L\| \|L(M_f - zI)h_n\|_1 \leq \|L\| \frac{1}{n}$, which implies $n \leq \|L\|$ for all $n$, a contradiction. So $S \subseteq \sigma(M_f)$.

Conversely, if $z \not\in S$, then there is $r > 0$ such that $m(f^{-1}(B(z, r))) = 0$. On $[0, 1] \setminus f^{-1}(B(z, r))$, define $g(x) = \frac{1}{f(x) - z}$ and on $f^{-1}(B(z, r))$, define $g(x) = 0$. Then $g$ is measurable on $[0, 1]$ and $\|g\|_\infty \leq \frac{1}{r}$. So $M_g(M_f - zI)(h) = h = (M_f - zI)g\phi(h)$ almost everywhere. Then $M_f - zI$ is invertible. Hence $\sigma(M_f) = S$.

$M_f$ may also be defined on $L^p[0, 1], 1 \leq p < \infty,$ by $M_f(g) = fg$. The norm and spectrum are the same as in the $L^1[0, 1]$ case. Finally, $M_f^* : L^q[0, 1] \to L^q[0, 1]$ is the same as $M_f$ because

$$\langle g, M_f^* h \rangle = \langle M_f(g), h \rangle = \int_{[0, 1]} (fg)h \, dm = \int_{[0, 1]} g(fh) \, dm = \langle g, fh \rangle.$$ 

Now we present two lemmas and the proofs of the two theorems above.

37
Lemma 1. For every $T \in L(X)$, $\sigma(T) \subseteq \overline{B(0, r)}$, where $r = \limsup_{n \to \infty} \|T^n\|^{1/n} < \infty$, i.e. $r(T) \leq r$. Also, for $|z| > r$, $(T - zI)^{-1} = -\sum_{n=0}^{\infty} z^{-n-1}T^n$.

Proof. Note $\|T^n\| \leq \|T\|^n$ implies that $r = \limsup_{n \to \infty} \|T^n\|^{1/n} \leq \|T\| < \infty$. For $|z| > r$, there is $\varepsilon > 0$ such that $|z| > r + \varepsilon$. By properties of limsup, we see that $\|T^n\|^{1/n} < r + \varepsilon$ for all except finitely many $n$. Hence, $z^{-n-1}T^n$ converges absolutely in $L(X)$. Thus the open ball of radius $\|z\|^{-1}$ centered at $z$ is in $\rho(T)$. So $T$ is invertible, i.e. $z \in \rho(T) = \mathbb{C} \setminus \sigma(T)$. Hence, $\sigma(T) \subseteq \overline{B(0, r)}$.

Lemma 2. Let $\Omega$ be a nonempty open subset of $\mathbb{C}$ contained in $\rho(T)$. If $f \in L(X)^*$, the function $g(z) = f((T - zI)^{-1})$ is holomorphic with derivative $g'(z) = f(R_z(T))$. Proof. This follows from the continuity of $f$ and the remark below the resolvent identity.

Proof of Gelfand’s Theorem. By lemma 1, $\sigma(T)$ is bounded in $\mathbb{C}$. Hence, $(T - zI)^{-1}$ is invertible. By the lemma on inverses, we get $T - wI = (T - zI) - (w - z)I$ is also invertible if $|w - z| < \|(T - zI)^{-1}\|^{-1}$. Thus the open ball of radius $\|(T - zI)^{-1}\|^{-1}$ centered at $z$ is in $\rho(T)$. So $\rho(T)$ is open and $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is closed.

Finally, we show $\sigma(T) \neq \emptyset$. Assume $\sigma(T) = \emptyset$. Let $\Omega = \rho(T) = \mathbb{C}$. By lemmas 1 and 2, for $|z| > \|T\| \geq r$,

$$|g(z)| \leq \|f\| \|(T - zI)^{-1}\| \leq \|f\| \sum_{n=0}^{\infty} |z|^{-n-1} \|T^n\| = \frac{\|f\|}{|z| - \|T\|} \to 0 \quad \text{as } z \to \infty.$$  

Hence, $g(z)$ is a bounded entire function. By Liouville’s theorem, $f((T - zI)^{-1}) = g(z) = 0$. Then $\|(T - zI)^{-1}\| = \sup\{|f((T - zI)^{-1})| : f \in L(X)^*, \|f\| \leq 1\} = 0$, which is absurd. So $\sigma(T) \neq \emptyset$.

Proof of the Gelfand-Mazur Theorem. Let $r = \limsup_{n \to \infty} \|T^n\|^{1/n}$. First we show there is a $z \in \sigma(T)$ with $|z| = r$. If $r = 0$, then $\emptyset \neq \sigma(T) \subseteq \overline{B(0, r)}$ implies $\sigma(T) = \{0\}$. Next we consider $r > 0$. Assume $\sigma(T) \cap \{z : |z| = r\} = \emptyset$. Take $R$ such that $r(T) = \max\{|z| : z \in \sigma(T)\} < R < r$. Then $\sigma(T) \subseteq \{z : |z| \leq r(T)\}$.

For all $f \in L(X)^*$, by lemmas 1 and 2, $g(z) = f((T - zI)^{-1}) = -\sum_{n=0}^{\infty} f(T^n)z^{-n-1}$ on $\{z : |z| > r\}$, hence also on $\{z : |z| = r(T)\}$ by the uniqueness of Laurent series on annulus. Then it converges absolutely on $|z| = R$. So sup\{|$f(T^n)$|$|R^n+1 : n = 0, 1, 2, \ldots| < \infty$. By the uniform boundedness principle, we get $c = \sup\{|T^n/R^n+1| : n = 0, 1, 2, \ldots| < \infty$. Hence, $\|T^n\| \leq cR^{n+1}$. Then $\|T^n\|^{1/n} \leq c^{1/n}R^{1+1/n}$. Taking limsup, we get $r \leq R$, a contradiction.

Next we show $r = \inf\{|T^n|^{1/m} : m = 1, 2, 3, \ldots|$. For positive integers $m, n$, we have $n = qm + k$ with $k = 0, 1, \ldots, m - 1$. Then $\|T^n\| \leq \|T^n\|^{q/m}|T|^k$. So $\|T^n\|^{1/n} \leq \|T^n\|^{q/m}|T|^k/n$. Fix $m$ and let $n \to \infty$, since $1 = m(q/n) + (k/n)$, we get $k/n \to 0$ and $q/n \to 1/m$. So $r = \limsup_{n \to \infty} \|T^n\|^{1/n} \leq \|T^n\|^{1/m}$. Taking infimum over $m$, we get the result $r \leq \inf\{|T^n|^{1/m} : m = 1, 2, 3, \ldots| \leq \liminf_{m \to \infty} |T^n|^{1/m} \leq \limsup_{m \to \infty} |T^n|^{1/m} = r$. \hfill \Box

§2. Projections and Complemented Subspaces. In the literature, vector subspaces are sometimes called linear manifolds. For convenience, below the term “subspaces” will mean closed vector subspaces of Banach spaces.
Definition. A subspace $E$ of a Banach space $X$ is complemented iff there is a subspace $F$ of $X$ such that $E \cap F = \{0\}$ and $E + F = X$. Such $F$ is called a complementary subspace for $E$. (In algebra, we write $X = E \oplus F$ and call it an internal direct sum.)

Remarks. (1) In the definition, if $x = y + z = y' + z'$ for $y, y' \in E$ and $z, z' \in F$, then $y - y' = z' - z \in E \cap F = \{0\}$ implies $y = y'$ and $z = z'$. So every $x$ has a unique representation as $y + z$ with $y \in E, z \in F$.

(2) We have $\dim F = \text{codim} E$ (i.e. $\dim X/E$) since if $B$ is a basis of $F$, then $\pi(B)$ is a basis of $X/E$, where $\pi : X \to X/E$ is the quotient map.

Examples. (1) If $\dim E = n < \infty$, then $E$ is complemented. (To see this, let $\{x_1, \ldots, x_n\}$ be a basis of $E$. By the Hahn-Banach theorem, for $i = 1, \ldots, n$, there is $f_i \in X^*$ such that $f_i(x_i) = 1$ and $f_i(x_j) = 0$ for $i \neq j$. Then it is easy to check that $F = \bigcap_{i=1}^n \ker f_i$ is a complementary subspace of $E$.)

(2) If $\text{codim} E < \infty$, then $E$ is complemented. (To see this, suppose $\dim(X/E) = n < \infty$. Let $\{x_1 + E, \ldots, x_n + E\}$ be a basis of $X/E$. Let $F$ be the linear span of $\{x_1, \ldots, x_n\}$. Then $\dim F = n < \infty$ implies $F$ is complete, hence closed. It is easy to check $E \cap F = \{0\}$ and $E + F = X$.)

(3) Every subspace $M$ in a Hilbert space $H$ is complemented by its orthogonal complement $M^\perp$, i.e. we have $H = M \oplus M^\perp$. (In 1971, Lindenstrauss and Tzafriri proved the converse, namely if every subspace of a Banach space is complemented, then the Banach space is isomorphic to a Hilbert space.)

(4) $e_0$ is uncomplemented in $\ell^\infty$. See [M], pp. 301-302.

(5) In $L^p = L^p(-\pi, \pi)$, let $H^p$ be the closed linear span of $e^{in\theta}$ ($n \geq 0$). M. Riesz proved that for $1 < p < \infty$, $H^p$ is complemented in $L^p$ by the closed linear span of $e^{in\theta}$ ($n < 0$). D. J. Newman proved that $H^1$ is uncomplemented in $L^1$. R. Arens and P. C. Curtis proved that $H^\infty$ is uncomplemented in $L^\infty$.

Definition. An operator $P \in L(X)$ is a projection iff $P^2 = P$, i.e. $P|_{\text{ran} P} = I|_{\text{ran} P}$.

Remarks. If $P$ is a projection, then $Q = I - P$ is a projection since $(I - P)^2 = I - 2P + P^2 = I - P$. Also, $\ker P = \text{ran}(I - P)$ since $Px = 0$ implies $x = x - Px = (I - P)x$ and conversely, $P((I - P)x) = Px - P^2 x = 0$. Similarly, $\text{ran } P = \text{ran}(I - Q) = \ker Q = \ker (I - P)$. So $P$ is always closed.

Theorem. If $P$ is a projection, then $\text{ran } P$ and $\ker P$ complement each other, i.e. $X = \text{ran } P \oplus \ker P$.

Proof. Since $\ker P = \text{ran}(I - P), x = Px + (I - P)x$ and $x \in \text{ran } P \cap (\ker P)$ implies $x = Px = 0$, we get $X = \text{ran } P \oplus \ker P$.

Theorem. A subspace $E$ of $X$ is complemented iff $E = \text{ran } P$ for some projection $P \in L(X)$.

Proof. The if direction follows from the last theorem. For the only-if direction, let $F$ be a complementary subspace of $E$. Then each $x \in X$ can be written as $x = y + z$ for some unique $y \in E$ and $z \in F$. Define $Px = y$. Then $P = E$ since $y = y + 0 \in E$ implies $Py = y$. Also, $P^2 x = Py = y = Px$, i.e. $P^2 = P$.

For continuity, consider the graph of $P$. If $(x_n, Px_n) \to (x, y)$, then writing $x_n = y_n + z_n$, where $y_n \in E$ and $z_n \in F$, we get $y_n = Px_n \to y \in E$ since $E$ is closed. So $z_n = x_n - y_n \to x - y \in F$ since $F$ is closed. Hence $x = y + (x - y)$ and by the definition of $P$, $y = Px$. By the closed graph theorem, $P$ is bounded.

Corollary. If $E$ is complemented in $X$, then $E^\perp$ is complemented in $X^\ast$.

Proof. Let $P \in L(X)$ be a projection with $\text{ran } P = E$, then $(P^\ast)^2 = P^\ast P^\ast = (PP)^\ast = P^\ast$, i.e. $P^\ast$ is a projection in $L(X^\ast)$ and $E^\perp = \text{ran } P^\perp = \ker P^\ast$ is closed and complemented (by $\text{ran } P^\ast$) from above.

Theorem. $T \in L(X,Y)$ is left-invertible (i.e. there is $S \in L(Y,X)$ such that $ST = I$) iff $T$ is injective and $\text{ran } T$ is closed and complemented.

Proof. For the if direction, let $P \in L(Y)$ be the projection onto $\text{ran } T$, then $T_0 = P \circ T : X \to \text{ran } T$ is bijective. Let $S = T_0^{-1} \circ P$, then $ST = I$. For the only-if direction, if $S \in L(Y, X)$ is such that $ST = I$, then
§3. **Compact Operators.** Finite rank operators (i.e. operators whose ranges are finite dimensional) are easy to understand by using linear algebra. In this section, we will study a class of operators related to the finite rank operators. First we recall the following facts:

1. For any normed vector space $V$, if the closed unit ball of $V$ is compact, then $\dim V < \infty$.

2. **(Metric Compactness Theorem)** In a metric space $M$, a set $S$ in $M$ is compact iff $S$ is sequentially compact (i.e. every sequence in $S$ has a convergent subsequence with limit in $S$) iff $S$ is complete and totally bounded (i.e. for every $\varepsilon > 0$, there are $x_1, \ldots, x_n \in S$ such that $B(x_1, \varepsilon) \cup \cdots \cup B(x_n, \varepsilon) \supseteq S$ and we say $\{x_1, \ldots, x_n\}$ is $\varepsilon$-dense in $S$). It is easy to check that a set is $\varepsilon/2$-dense implies its closure is $\varepsilon$-dense. Hence, a set is totally bounded if and only if its closure is totally bounded.

3. **(Arzela-Ascoli Theorem)** For a compact set $M$, a set $S$ in $C(M, \mathbb{K})$ is (sequentially) compact iff $S$ is closed, bounded and equicontinuous in $C(M, \mathbb{K})$, where equicontinuity means for every $\varepsilon > 0$, there is a $\delta > 0$ such that for all $f \in S$ and for all $x, y \in M$, $d(x, y) < \delta$ implies $|f(x) - f(y)| < \varepsilon$.

**Definition.** Let $X, Y$ be Banach spaces and $B$ be the open unit ball of $X$. A linear function $K : X \to Y$ is **compact** iff $K(B)$ is precompact, i.e. $\overline{K(B)}$ is compact, in $Y$. (By the metric compactness theorem, this is equivalent to the condition that for every bounded sequence $\{x_n\}$ in $X$, the sequence $\{K(x_n)\}$ has a convergent subsequence in $Y$ or to $K(B)$ is totally bounded).

**Remark.** Since $\overline{K(B)}$ is compact, it cannot contain any closed ball (which is never compact) in infinite dimensional spaces. So compact operators are considered “small” operators.

**Theorem (Properties of Compact Operators).** Let $X, Y, Z$ be Banach spaces.

(a) Finite rank operators $F \in L(X, Y)$ (i.e. $\dim \ker F < \infty$) are compact. If $K \in L(X, Y)$ is compact, then $\ker K$ contains no infinite dimensional closed subspaces. In particular, if $\ker K$ is also closed, then $K$ has finite rank.

(b) If $K_1, K_2$ are compact and $c \in \mathbb{C}$, then $K_1 + cK_2$ is compact.

(c) If $K \in L(X, Y)$ is compact and $T \in L(Y, Z)$, then $TK$ is compact.

(d) If $K \in L(Y, Z)$ is compact and $T \in L(X, Y)$, then $KT$ is compact.

(e) If $K \in L(Y, Z)$ is compact and invertible, then $\dim X = \dim Y < \infty$.

(f) The restriction $K|_V$ of a compact operator $K \in L(X, Y)$ to a closed subspace $V$ of $X$ is compact.

(g) If $K \in L(X, Y)$ is compact, then $\ker K$ is separable.

(h) If for $n = 1, 2, 3, \ldots$, $K_n \in L(X, Y)$ is compact and $K_n$ converges to $K$, then $K$ is compact.

(i) $K \in L(X, Y)$ is compact iff $K^* \in L(Y^*, X^*)$ is compact.

**Remarks.** (1) In the case $X = Y = Z$, parts (b), (c), (d), (h) imply the set of all compact operators is a closed two-sided ideal of $L(X)$. (2) Part (i) of the theorem is called Schauder’s theorem in some literatures.

**Examples.** (1) Let $X = Y = \ell^p$ ($1 \leq p \leq \infty$). For $a = (a_1, a_2, a_3, \ldots) \in c_0$, define $K(x_1, x_2, x_3, \ldots) = (a_1x_1, a_2x_2, a_3x_3, \ldots)$. Then $\|K\| \leq \|a\|_\infty$. Next, define $K_n(x_1, x_2, x_3, \ldots) = (a_1x_1, a_2x_2, \ldots, a_nx_n, 0, 0, \ldots)$,
which is finite rank, hence compact. Then \( \|K - K_n\| \leq \sup\{|a_j| : j > n\} \to 0 \) as \( \limsup_{n \to \infty} |a_n| = 0 \).

By property (h), \( K \) is compact.

(2) Let \( X = Y = C([0,1]) \) and \( G \in C([0,1]^2) \). Define \( (Kf)(x) = \int_0^1 G(x,y)f(y) \, dy \). This is called the Fredholm integral operator. Note that \( K \in L(X) \) and \( \|K\| \leq \|G\|_\infty \). If \( G(x,y) = F(x)H(y) \) for some \( F, H \in C([0,1]) \), then \( K \) has at most rank 1. Similarly, if \( G(x,y) = \sum_{j=1}^n F_j(x)H_j(y) \), then \( K \) has finite rank.

By the Stone-Weierstrass theorem, we can approximate \( G \) uniformly by functions of the form \( \sum_{j=1}^n F_j(x)H_j(y) \).

So we can approximate \( K \) by finite rank operators. Therefore, \( K \) is compact.

(3) Let \( X = Y = L^2([0,1]) \) and \( G \in L^2([0,1]^2) \). Define \( K \) as above. Then \( K \in L(X) \) and \( \|K\| \leq \|G\|_2 \). By the reasoning above, \( K \) is compact (as continuous functions are dense in \( L^2 \)) by (h).

(4) Let \( X = C^1([0,1]) \) be the set of functions with continuous derivatives on \([0,1] \). For \( f \in C^1([0,1]) \), let \( \|f\|_{C^1([0,1])} = \|f\|_\infty + \|f'\|_\infty \). This is a complete norm by properties of uniform convergence. So \( C^1([0,1]) \) is a Banach space. Let \( Y = C([0,1]) \) and \( K : X \to Y \) be the inclusion map \( K(f) = f \). Then \( K \) is compact by the Arzela-Ascoli theorem because \( \|f_n\|_{C^1([0,1])} \leq 1 \) implies \( \|f_n\|_\infty \leq 1 \) (hence \( \{f_n\} \) bounded in \( C([0,1]) \)) and \( \|f_n'\|_\infty \leq 1 \) (hence \( \{f_n\} \) is equicontinuous in \( C([0,1]) \)) by the mean-value theorem.

**Proof of Properties of Compact Operators.** Let \( B \) and \( B' \) denote the open unit balls of \( X \) and \( Y \) respectively.

(a) For the first statement, \( \overline{F(B)} \subseteq (\|F\|B') \cap (\text{ran } F) \) and \( \text{dim ran } F < \infty \) imply \( (\|F\|B') \cap (\text{ran } F) \) is compact and hence \( F(B) \) is compact. For the second statement, let \( Z \) be a closed subspace of \( \text{ran } K \), then \( W = K^{-1}(Z) \) is closed in \( X \). Consider the surjection \( K|_W : W \to Z \). By the open mapping theorem, \( K|_W \) sends the open unit ball \( B_W \) of \( W \) to an open neighborhood \( K(B_W) \) of \( 0 \) in \( Z \). Then \( K(B_W) \) is a compact neighborhood of \( 0 \) in \( Z \). This implies \( Z \) is finite dimensional.

(b) \( K_1 + cK_2 \) compact follows from \( (K_1 + cK_2)(B) \subseteq K_1(B) + cK_2(B) \), which is compact as it is the image of \( K_1(B) \times K_2(B) \) under the continuous function \( g(x,y) = x + cy \).

(c) \( TK \) compact follows from \( TK(B) \subseteq T(\overline{K(B)}) \), which is compact.

(d) \( KT \) compact follows from \( KT(B) \subseteq K(\|T\|B') \subseteq T(\overline{K(B')}) \), which is compact.

(e) By (c) and (d), \( K^{-1}K = I \) and \( KK^{-1} = I \) are compact and hence the closed unit balls of \( X \) and \( Y \) are compact. Then \( X, Y \) are finite dimensional. \( K \) invertible implies the dimensions are the same.

(f) \( K|_V \) compact follows from \( K|_V(B \cap V) \subseteq \overline{K(B)} \), which is compact.

(g) This follows from \( K(B) \) totally bounded, hence separable, and \( \text{ran } K = \bigcup_{n=1}^{\infty} nK(B) \).

(h) Let \( B \) be the open unit ball of \( X \). To show \( \overline{K(B)} \) compact, it is enough to show \( K(B) \) is totally bounded. For \( \varepsilon > 0 \), take \( n \) with \( \|K_n - K\| < \varepsilon/3 \). Since \( K_n(B) \) is compact, it is totally bounded. So there is a finite set \( \{x_1, \ldots, x_m\} \subseteq B \) such that \( \{K_n(x_1), \ldots, K_n(x_m)\} \) is \( (\varepsilon/3) \)-dense in \( K_n(B) \). Hence, for every \( y \in B \), there is \( j \) with \( \|K_n(y) - K_n(x_j)\| < \varepsilon/3 \), so

\[ \|K(y) - K(x_j)\| \leq \|K(y) - K_n(y)\| + \|K_n(y) - K_n(x_j)\| + \|K_n(x_j) - K(x_j)\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \]

Hence, \( \{K(x_1), \ldots, K(x_m)\} \) is \( \varepsilon \)-dense in \( K(B) \). Therefore, \( K(B) \) is totally bounded.

(i) If \( K \) is compact, then let \( \{y_n\} \) be a sequence in \( Y^* \) with \( \|y_n\| \leq 1 \). Let \( U \) be the closed unit ball of \( X \), then \( \overline{K(U)} \) is compact in \( Y \). Since for every \( x, z \in K(U) \), \( \|y_n(x) - y_n(z)\| \leq \|y_n\||x - z| \leq ||x - z|| \), the functions \( y_n \) are equicontinuous in \( C(K(U), \mathbb{C}) \). By the Arzela-Ascoli theorem, there is a subsequence \( \{y_{n_k}\} \)

41
convergent in $C(K(U), \mathbb{C})$. Since $K^*y_n = y_n \circ K$, the sequence $\{K^*y_n\}$ converges uniformly on $U$. Since norm of $T$ in $X^*$ is sup-norm of $T$ on $U$, $K^*y_n$ converges in $X^*$. Hence $K^*$ is compact.

Conversely, $K^*$ compact implies $K^{**}$ is compact, which implies $K = K^{**}|_X$ is compact. 

From property (h), we know the limit of finite rank operators is compact. This raised the question of whether compact operators are always limit of finite rank operators or not. In the case $Y = X$ is a Hilbert space, it is true and will be shown in the next chapter. Below we will prove it for a separable Hilbert space with the help of the following theorem.

**Theorem.** If $K \in L(X,Y)$ is compact, then $K$ is completely continuous, which means that for every $\{x_n\}$ $w$-converges to $x$ in $X$, $Kx_n$ norm-converges to $Kx$ in $Y$. For reflexive $X$, the converse is true.

**Proof.** For the first statement, assume $Kx_n$ does not converge to $Kx$. Then there are $\varepsilon > 0$ and subsequence $\{x_{n_k}\}$ such that $\|Kx_{n_k} - Kx\| \geq \varepsilon$. Since $\{x_{n_k}\}$ $w$-converges to $x$, by the uniform boundedness principle, $\{x_{n_k}\}$ is bounded. By compactness of $K$, there is a subsequence $x_{n_{k_j}}$ such that $Kx_{n_{k_j}}$ norm-converges (hence also $w$-converges) to some $z$. Since $\|z - Kx\| = \lim_{j \to \infty} \|Kx_{n_{k_j}} - Kx\| \geq \varepsilon$, $z \neq Kx$. Since $x_n \to^w x$, for every $f \in Y^*$, we have $K^*(f) \in X^*$ and $f(Kx_{n_{k_j}} - Kx) = K^*(f)(x_{n_{k_j}} - x) \to 0$, i.e. $Kx_{n_{k_j}}$ $w$-converges to $Kx$. This leads to $Kx = z$, a contradiction.

For the second statement, since $X$ is reflexive, for every bounded sequence $\{x_n\}$ in $X$, by the Eberlein-Smulian theorem, there is a subsequence $\{x_{n_k}\}$ $w$-converges to some $w$. Then $\{Kx_{n_k}\}$ converges to $Kw$ by assumption. Therefore, $K$ is compact. 

**Theorem.** Let $H$ be a separable Hilbert space and $K \in L(H)$ be a compact operator. Then $K$ is the limit of a sequence of finite rank operators in $L(H)$ under the norm topology.

**Proof.** For $K$ with finite rank, take every term to be $K$. For compact $K$, not finite rank, by property (g) of compact operators, ran $K$ is separable. Let $\{y_1, y_2, y_3, \ldots\}$ be an orthonormal basis of ran $K$ and $P_n x = \sum_{j=1}^n \langle x, y_j \rangle y_j$ be the projection onto the closed linear span of $\{y_1, \ldots, y_n\}$. By Bessel's inequality, $\|P_n\| = 1 = \|I - P_n\|$. For $1 \leq m \leq n$, we have $P_n P_m = P_m$ and so $(I - P_n)(I - P_m) = I - P_n - P_m + P_n P_m = I - P_n$. Then 

$$\|K - P_n K\| = \|(I - P_n)K\| = \|(I - P_n)(I - P_n)K\| \leq \|(I - P_n)K\| = \|K - P_n K\|.$$ 

Hence $\{\|K - P_n K\|\}$ is a decreasing sequence of nonnegative numbers. Assume its limit is $\eta \neq 0$. Then for every $n$, there is $x_n \in H$ such that $\|x_n\| = 1$ and $\|(I - P_n)Kx_n\| > \eta/2$. By the Eberlein-Smulian theorem, since Hilbert spaces are reflexive, $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ converges weakly to some $x$. By the last theorem, $Kx_{n_k}$ converges in norm to $u = Kx$. Now $P_n u$ converges to $u$ in norm. Then 

$$\frac{\eta}{2} < \|(I - P_n)Kx_{n_k}\| \leq \|(I - P_n)(Kx_{n_k} - u)\| + \|(I - P_n)u\| \leq \|Kx_{n_k} - u\| + \|u - P_n u\| \to 0,$$

a contradiction. Therefore, $\|K - P_n K\| \to \eta = 0$ and $P_n K$ is finite rank.

**Definition.** (1) A Banach space $Y$ has the approximation property if for every Banach space $X$, every compact operator in $L(X,Y)$ is the limit of a sequence of finite rank operators in $L(X,Y)$.

(2) A sequence $\{x_n\}$ in a Banach space $Y$ is a Schauder basis of $Y$ iff for every $y \in Y$, there is a unique sequence $\{c_n\}$ of scalars such that $y = \sum_{n=1}^\infty c_n x_n$. (Such spaces are clearly separable.)

**Remarks.** It is known that if $Y$ has a Schauder basis, then $Y$ has the approximation property (see [M], p. 364) and in particular, every compact operator in $L(Y)$ is the limit of a sequence of finite rank operators in $L(Y)$. See [CL], pp. 212-213. In 1932, Banach conjectured that every Banach space $Y$ has the approximation property and further conjectured that every separable Banach space has a Schauder basis. On November
6, 1936, Mazur offered a goose as a prize for the solution of these problems in (problem 153 of) the famous “Scottish book” of open problems kept at the Scottish Coffee House in Lwów, Poland by Banach, Mazur, Ulam and other mathematicians.

In 1955, A. Grothendieck proved that $Y$ has the approximation property iff for every compact subset $W$ of $Y$ and every $\varepsilon > 0$, there is a finite rank operator $T \in L(Y)$ such that for all $y \in W$, $\|Ty - y\| < \varepsilon$. Thus to check the approximation property, there is no need to involve other Banach spaces $X$. Separable Hilbert spaces, $c_0$ and $\ell^p (1 \leq p < \infty)$ have the approximation property.

Finally, in 1971, Swedish mathematician and pianist Per Enflo showed that there is a separable reflexive Banach space $Y$ and a compact operator in $L(Y)$ that is not the limit of any sequence of finite rank operators in $L(Y)$. This refuted both conjectures. About a year after solving the problem, Enflo traveled to Warsaw to give a lecture on his solution, after which he was awarded the goose. Enflo’s solution was published in Acta Mathematica, vol. 130 (1973), pp. 309-317.

Next we will look at theorems about compact operators, which are useful for differential equations.

**Lemma.** If $K \in L(X)$ is compact and $c \neq 0$, then $N = \ker(K-cI)$ is finite dimensional and $M = \text{ran}(K-cI)$ is closed and finite codimensional.

**Proof.** For $N$, note that $|K|_N$ is compact and $|K|_N = cI$, which is invertible and hence $N$ is finite dimensional by property (e). For $M$, note $M^\perp = \ker(K^* - cI)$, which is finite dimensional. If we can show $M$ is closed, then $(X/M)^* = M^\perp$ is finite dimensional and since $\text{dim}(X/M) = \text{dim}(X/M)^*$, $M$ must be finite codimensional.

Let $Z$ be a complementary subspace of $N = \ker(K-cI)$. Since $Z \cap N = \{0\}$, $S = (K-cI)|_Z : Z \to X$ is injective. To show $M$ is closed, since $M = \text{ran}S$, by the lower bound theorem, it suffices to show $S$ is bounded below.

Assume $S$ is not bounded below. Then there is $z_n \in Z$, $\|z_n\| = 1$ and $S(z_n) \to 0$. Since $K$ is compact, passing to a subsequence, we may assume $K(z_n) \to w$. Then $z_n = (K-S)(z_n)/c \to w/c$, which is in $Z$ as $Z$ is closed. As $\|z_n\| = 1$, so $\|w\| = |c| \neq 0$. Also, $K(z_n) \to K(w/c)$. By the uniqueness of limit, $w = K(w/c)$. Then $w \in \ker(K-cI) \cap Z = \{0\}$, contradicting $\|w\| \neq 0$.

**Theorem (Riesz-Fredholm).** Let $K \in L(X)$ be compact, $c \neq 0$, $N_i = \ker(K-cI)^i$ and $M_i = \text{ran}(K-cI)^i$.

(a) $K(N_i) \subseteq N_i$ and $\text{dim} N_i < \infty$. $N_1 \subseteq N_2 \subseteq N_3 \subseteq \cdots$ and there is a least $j$ such that $N_j = N_{j+1} = N_{j+2} = \cdots$.

(b) $K(M_i) \subseteq M_i$, $M_i$ is closed and codim $M_i < \infty$. $M_1 \supseteq M_2 \supseteq M_3 \supseteq \cdots$ and there is a least $k$ such that $M_k = M_{k+1} = M_{k+2} = \cdots$.

(c) $j = k$ and $X = M_j \oplus N_j$. Also, $(K-cI)|_{M_j} \in L(M_j)$ is invertible and $(K-cI)|_{N_j} \in L(N_j)$ is nilpotent of index $j$ (i.e. $(K-cI)|_{N_j}^{j-1} \neq 0$, but $(K-cI)|_{N_j}^j \equiv 0$.)

(d) $\text{dim} \ker(K-cI) = \text{codim} \text{ran}(K-cI) = \text{dim} \ker(K^* - cI) = \text{codim} \text{ran}(K^* - cI) < \infty$. In particular, $K-cI$ is injective iff $K^* - cI$ is injective iff $K^* - cI$ is surjective.

**Proof.** (a) Observe that $z \in N_i$ implies $(K-cI)^i(Kz) = K(K-cI)^i(z) = 0$ (i.e. $Kz \in N_i$). So $K(N_i) \subseteq N_i$.

Next, $K$ compact implies $T = (K-cI)^i = (-c)^iI$ compact. So $N_i = \ker(K-cI)^i = \ker(T + (-c)^iI)$ is finite dimensional by the lemma.

Now, $N_i \subseteq N_{i+1}$ because $(K-cI)^i(x) = 0$ implies $(K-cI)^{i+1}(x) = 0$. Assume $N_i \subseteq N_{i+1}$ for all $i$. Pick $x_i \in N_i$ with $\|x_i\| \leq 2$ and $\|x_i + N_{i-1}\| = 1$. (This is possible by taking $x + N_{i-1} \in N_i/N_{i-1}$ with $\|x + N_{i-1}\| = 1$, then there is $y \in N_{i-1}$ such that $\|x + y\| \leq 2$ and we can let $x_i = x + y$, then $x_i + N_{i-1} = x + N_{i-1}$.) If $i < j$, then $x_i \in N_i$ implies $Kx_i \in N_i$ and

$$Kx_j - Kx_i = cx_j + (Kx_j - cx_j) - Kx_i \in cx_j + N_{j-1} + N_i = cx_j + N_{j-1} = c(x_j + N_{j-1}).$$
So  \( \|Kx_j - Kx_i\| \geq |c| > 0 \). Then  \( \{Kx_i\} \) has no convergent subsequence, contradicting  \( K \) is compact. Therefore, there is a least  \( j \) such that  \( N_j = N_{j+1} \). Since  \( x \in N_{j+1} \) implies  \((K-cI)x \in N_{j+1} = N_j\), which implies  \( x \in N_{j+1} \), so  \( N_{j+1} = N_{j+2} \) and so on.

(b) is similar to (a).

(c) To show  \( j = k \), suppose  \( a \in N_{k+1} \), i.e.  \((K-cI)^{k+1}(a) = 0\). Take  \( m > 0 \) such that  \( m + k \geq j \). Since  \((K-cI)^{b}(a) \in M_k = M_{m+k} \), we have  \((K-cI)^{b}(a) = (K-cI)^{m+k}(b) \) for some  \( b \in X \). Since  \( N_j = \cdots = N_{m+k} = N_{m+k+1} \), so  \((K-cI)^{j+1}(a) = (K-cI)^{m+k+1}(b) = (K-cI)^{m+k}(b) = (K-cI)^{b}(a)\). So  \( N_{k+1} = N_k \). By minimality of  \( j \), we get  \( j \leq k \).

For the converse, note that  \( N_j^\perp = \overline{\text{ran}(K^* - cI)}^{w*} = \overline{\text{ran}(K^* - cI)^j} = \text{ran}(K^* - cI)^j \) by the closed range theorem and  \( M_j^\perp = \text{ker}(K^* - cI)^j \). So applying the same reasoning to  \( K^* \), we get  \( k \leq j \). Therefore  \( j = k \).

Next, we show  \( X = M_j \oplus N_j \). Let  \( x \in X \). Since  \((K-cI)^j(x) \in M_j = M_{2j} \),  \((K-cI)^j(x) = (K-cI)^{2j}(y) \) for some  \( y \in X \). Write  \( x = (K-cI)^j(y) + z \). Then  \((K-cI)^j(z) = 0 \), i.e.  \( z \in N_j \). So  \( X = M_j \oplus N_j \). Now for  \( x \in M_j \cap N_j \), there is  \( y \in X \) such that  \( x = (K-cI)^j(y) \) and  \( 0 = (K-cI)^j(x) = (K-cI)^{2j}(y) \). Since  \( N_{2j} = N_j \), we have  \( x = (K-cI)^j(y) = 0 \). Therefore,  \( X = M_j \oplus N_j \).

Next we show  \((K-cI)|_{M_j} : M_j \to M_j \) is injective and surjective. For  \( x \in \text{ker}(K-cI)|_{M_j} \), there is  \( y \) such that  \( x = (K-cI)^j(y) \in M_j \) and  \((K-cI)x = 0 \). Then  \( y \in N_{j+1} \) so that  \( x = (K-cI)^j(y) = 0 \). Hence,  \((K-cI)|_{M_j} \) is injective. Also, for  \( z \in M_j = M_{j+1} \), we have  \( z = (K-cI)^j+1(w) = (K-cI)(K-cI)^j(w) \) for some  \( w \) and so  \( z \in \text{ran}(K-cI)|_{M_j} \). Hence,  \((K-cI)|_{M_j} \) is surjective. Therefore,  \((K-cI)|_{M_j} \) is invertible.

Finally, since  \( N_{j-1} \subset N_j \), there is  \( x \in N_j \setminus N_{j-1} \). So  \((K-cI)|_{N_j}^{j-1}(x) \neq 0 \). By definition of  \( N_j \),  \((K-cI)|_{N_j}^{j} \equiv 0 \). So  \((K-cI)|_{N_j} \) is nilpotent of index  \( j \).

(d) By (c),  \( X = M_j \oplus N_j \) and  \((K-cI)|_{M_j} \) is invertible. By (a),  \( \dim N_j < \infty \). Hence,  
\[
\dim \ker(K-cI) = \dim \ker(K-cI)|_{N_j} = \text{codim ran}(K-cI)|_{N_j} = \text{codim ran}(K-cI) < \infty.
\]

Similarly,  \( \dim \ker(K^* - cI) = \text{codim ran}(K^* - cI) < \infty \). For the middle equality, by the kernel-range relations and the duality theorem,
\[
\ker(K^* - cI) = (\ker(K-cI)^*) = (X/\text{ran}(K-cI))^* = \dim(X/\text{ran}(K-cI)) = \dim(X/\text{ran}(K-cI))^* = \dim \ker(K^* - cI).
\]

The following theorem of Riesz and Schauder on the spectrums of compact operators together with the Riesz-Fredholm theorem provided our understanding to the Sturm-Liouville boundary value problems.

**Theorem (Riesz-Schauder).** Let  \( K \in L(X) \) be compact.

(a) If  \( \dim X = \infty \), then  \( 0 \in \sigma(K) \). If  \( c \in \sigma(K) \) and  \( c \neq 0 \), then  \( c \) is an eigenvalue of  \( K \) and  \( K^* \) of finite multiplicities (i.e. the dimensions of the spaces of eigenvectors are finite).

(b)  \( \sigma(K) \) is a countable compact set and  \( 0 \) is the only possible limit point of  \( \sigma(K) \).

**Proof.** (a) By property (e) of compact operators, if a compact operator  \( K \in L(X) \) is invertible, then  \( \dim X < \infty \). The contrapositive asserts that  \( \dim X = \infty \) implies  \( 0 \in \sigma(K) \). For  \( c \in \sigma(K) \setminus \{0\} \),  \( K-cI \) is either not injective or not subjective. By part (d) of the Riesz-Fredholm theorem,
\[
0 < \dim \ker(K-cI) = \text{codim ran}(K-cI) = \dim \ker(K^* - cI) = \text{codim ran}(K^* - cI) < \infty.
\]

Therefore,  \( c \) is an eigenvalue of  \( K \) and  \( K^* \) of finite multiplicities.

(b) For  \( c \in \sigma(K) \setminus \{0\} \), by part (c) of the Riesz-Fredholm theorem,  \( A = (K-cI)|_{M_j} \) is invertible. By the lemma on inverses, for  \( |z-c| < \|A^{-1}\|^{-1} \), we know  \( (K-zI)|_{M_j} = A - (z-c)I \) is invertible.
Also, by part (c) of the Riesz-Fredholm theorem, 
\( T = (K - cI)|_{N_j} \) is nilpotent of index \( j \), i.e. \( T^j \equiv 0 \). 
By the Gelfand-Mazur theorem, \( \sigma(T) = \{0\} \) or observe that for \( \alpha \neq 0 \),
\[
(T - \alpha I)^{-1} = -\alpha^{-j}(T^{-j-1} + \alpha T^{-j-2} + \cdots + \alpha^{-1} I).
\]
Then, for \( z \neq c \), \( (K - zI)|_{N_j} = T - (z - c)I \) is invertible. So for \( 0 < |z - c| < \|A^{-1}\|^{-1} \), \( K - zI \) is invertible on \( X = M_j \oplus N_j \), i.e. \( z \notin \sigma(K) \). Hence \( c \) is an isolated point in \( \sigma(K) \). For \( n = 1, 2, 3, \ldots \), the set \( S_n = \sigma(K) \cap \{z : |z| \geq 1/n\} \) is finite (otherwise, by the Bolzano-Weierstrass theorem, \( S_n \) has a limit point \( c \), which cannot be isolated). Therefore, \( \sigma(K) = S_1 \cup S_2 \cup S_3 \cup \cdots \) is countable and 0 is the only possible limit point of \( \sigma(K) \).

In the beginning of the twentieth century, Fredholm inspired many mathematicians to investigate integral equations. These works led to the solutions of the Neumann and Dirichlet problems by single and double layer potential methods (see Folland’s *Intro. to PDE*, Chapter 3). The integral equations were mostly of the form
\[
\int_a^b G(s,t) x(t) \, dt - cx(s) = y(s).
\]
In case \( G \) and \( x \) are continuous, the first term on the left is a compact operator. The studies in these equations led to the theory of compact operators. The following were the results obtained for these equations.

**Corollary (Fredholm Alternatives).** Let \( X \) be a Banach space, \( K \in L(X) \) be compact and \( c \neq 0 \). Either
(a) \( K - cI \) is invertible or (b) \( 0 < \dim \ker(K - cI) < \infty \).

If (a) holds, then \( K^* - CI \) is invertible. If (b) holds, then \( 0 < \dim \ker(K - cI) = \dim \ker(K^* - cI) < \infty \).

Furthermore, \( (K - cI)x = y \) if and only if \( y \in \ker(K^* - cI) \). Also, \( (K^* - cI)x = y^* \) if and only if \( y^* \in \ker(K - cI) \).

**Proof.** By part (d) of the Riesz-Fredholm theorem, \( 0 \leq \dim \ker(K - cI) = \text{codim ran}(K - cI) < \infty \).

Alternative (a) is the case \( 0 = \dim \ker(K - cI) = \text{codim ran}(K - cI) \). Alternative (b) is the case \( 0 < \dim \ker(K - cI) < \infty \).

If (a) holds, then \( 0 = \dim \ker(K^* - cI) = \text{codim ran}(K^* - cI) \). If (b) holds, then \( 0 < \dim \ker(K - cI) = \dim \ker(K^* - cI) < \infty \).

The furthermore statement follows as \( \text{ran}(K - cI) = \overline{\text{ran}(K - cI)} = \frac{1}{2} \ker(K^* - cI) \) and \( \text{ran}(K^* - cI) = \overline{\text{ran}(K^* - cI)^{\perp}} = \ker(K - cI)^{\perp} \) by using the closed range theorem and the kernel-range relations.

In ordinary differential equation, the Sturm-Liouville boundary value problems (see Boyce and DiPrima’s *Elementary Differential Equations and Boundary Value Problems*, Chapter 11) are important. It is well-known that the corresponding Sturm-Liouville operators have real eigenvalue sequence tending to infinity. Being unbounded operators, when they are injective, it is known (see Gohberg, Goldberg and Kaashoek’s *Basic Classes of Linear Operator*, Chapter 6) to have inverses, which are compact integral operators.

One of the most important problems in operator theory is to determine if every operator \( T \in L(X) \) has a nontrivial closed invariant subspace \( M \) (i.e. \( \{0\} \subset M \subset X \) and \( T(M) \subseteq M \)). For \( X = \ell^2 \), Enflo proved that there exists operators without nontrivial closed invariant subspaces. The case \( X \) is a Hilbert space is still open. For compact operators, not only do they have nontrivial closed invariant subspaces, but we also have the following stronger results.

**Lomonosov’s Theorem.** Let \( X \) be an infinite dimensional Banach space over \( \mathbb{C} \) and \( K \) be a nonzero compact operator. Then there exists a closed subspace \( M \) of \( X \) such that \( \{0\} \subset M \subset X \) and for every \( S \in L(X) \) commuting with \( K \) (i.e. satisfying \( SK = KS \)), we have \( S(M) \subseteq M \). Such a closed subspace \( M \) is called a nontrivial hyperinvariant subspace of \( K \).

**Proof.** (Due to H. M. Hilden) Let \( \Gamma = \{S \in L(X) : SK = KS\} \), which is called the commutant of \( K \). For every \( y \in X \), \( \Gamma_y = \{Sy : S \in \Gamma\} \) is a closed subspace of \( X \) which contains \( I(y) = y \). If \( y \neq 0 \), then \( \{0\} \subset \Gamma_y \).
Also $S_0(\Gamma_y) \subseteq \Gamma_y$ for every $S_0 \in \Gamma$. So if there is a $y \neq 0$ such that $\Gamma_y \subset X$, then $M = \Gamma_y$ is a nontrivial hyperinvariant subspace of $K$.

In the case $\Gamma_y = X$ for all $y \neq 0$. Pick $x_0 \in X$ so that $Kx_0 \neq 0$, then $x_0 \neq 0$. Since $K$ is bounded, there is an open ball $B$ with center at $x_0$ inside $B(x_0, \|x_0\|/2) \cap K^{-1}(B(Kx_0, \|Kx_0\|/2))$. So for all $x \in B$, $\|x\| \geq \|x - x_0\| \geq \|x_0\|/2 > 0$ and $\|Kx\| \geq \|Kx_0 - Kx - Kx_0\| \geq \|Kx_0\|/2 > 0$ for all $x \in B$. Then $0 \notin B$ and $0 \notin K(B)$.

For every $y \in K(B)$, since $\Gamma_y = X$, there is some $S_y \in \Gamma$ such that $S_y(y) \in B$. Then $W_y = S^{-1}_y(B)$ is open. Since $\{W_y : y \in K(B)\}$ covers $K(B)$, there are $W_1, \ldots, W_n$ such that $K(B) \subseteq W_1 \cup \cdots \cup W_n$. Let $S_i = S^{-1}_y(B)$, then $S_i(W_i) \subseteq B \subseteq B(x_0, \|x_0\|/2)$ and $0 \notin B$ imply $S_i \neq 0$. So $d = \max\{\|S_1\|, \ldots, \|S_n\|\} > 0$.

Since $Kx_0 \in K(B)$, there are $S_i$ and $W_i$ such that $Kx_0 \in W_i$. Then $S_i, Kx_0 \in S_i(W_i) \subseteq B$ and $Kx_0 \in K(B)$. So there are $S_i$ and $W_i$ such that $Kx_0 \in W_i$ so that $S_i, Kx_0 \in B$. By an induction argument, for every positive integer $j$, there is $x_j = S_1j \cdots S_1Kx_0 = S_1j \cdots S_1K^jx_0 \in B$. Hence, $d^j\|K\|\|x_0\| \geq \|x_j\| \geq \|x_0\| - \|x_j - x_0\| \geq \|x_0\|/2$ and so $r(K) = \lim_{j \to \infty} \|K^j\|^{1/j} \geq 1/d > 0$. Then $\sigma(K)$ contains some $c \neq 0$.

By the Riesz-Schauder theorem, $c$ is an eigenvalue of $K$. Then $M = \ker(K - cl) = \{v \in X : Kv = cv\}$ is finite dimensional. Hence, $M$ is a closed subspace satisfying $\{0\} \subset M \subset X$. For every $S \in \Gamma$ and $v \in M$, we have $KSv = SKv = S(cv) = cSv$, which implies $S(M) \subseteq M$. So, $M$ is hyperinvariant.

**Remark.** In fact, Lomonosov proved a even stronger result, namely if $A \neq 0$ commutes with $B \neq 0$, which commutes with a nonzero compact operator, then $A$ has a nontrivial closed invariant subspace.

---

§4. Fredholm Operators. In this section, we study a special class of operators, for which we can associate an index that has deep connections with elliptic differential operators on manifolds. In the 1960s, Atiyah, Singer and Conner connected this analytic index on differential operators with a topological index on a manifold that generalized the winding number of a closed curve (i.e. a one dimensional manifold) around a point. This so-called Atiyah-Singer index theorem was a great achievement in the 20th century mathematics. We recommend Booss and Bleeker’s book *Topology and Analysis* for an understanding of this theorem.

**Definitions.** For Banach spaces $X$ and $Y$, $T \in L(X,Y)$ is a Fredholm operator if $(\text{ran} T)$ is closed, $\dim \ker T < \infty$ and $\text{codim} \text{ran} T < \infty$. For a Fredholm operator, the *index* of $T$ is $\text{ind} T = \dim \ker T - \text{codim} \text{ran} T$. In some literatures, the *cokernel* of $T$ is defined to be $\text{coker} T = Y/(\text{ran} T)$ and in that case, $\text{ind} T = \dim \ker T - \dim \text{coker} T$.

**Remarks.** If $Y/\text{ran} T$ is finite dimensional as a vector space, then $\text{ran} T$ is closed. To see this, let $W$ be a finite dimensional vector subspace of $Y$ such that $\text{ran} T \cap W = \{0\}$ and $\text{ran} T + W = Y$. Then $X/(\ker T) \oplus W$ is a Banach space. Define $f : (X/\ker T) \oplus W \to Y$ by $f([x],w) = Tx + w$. Since $T : X/\ker T \to \text{ran} T$ is an isomorphism, $f = T \oplus I$ is bijective and continuous. Hence $f$ is bounded below. Hence $\text{ran} f = T((X/\ker T) \oplus \{0\})$ is bounded below. Since $(X/\ker T) \oplus \{0\}$ is complete, $\text{ran} T = f((X/\ker T) \oplus \{0\})$ is complete, hence closed.

**Examples.** (1) If $T : X \to Y$ is invertible, then $T$ is Fredholm with $\ker T = \{0\}$, $\text{ran} T = Y$ and so $\text{ind} T = 0$.

(2) If $T_0 : X_0 \to Y_0$ and $T_1 : X_1 \to Y_1$ are Fredholm, then $T_0 \oplus T_1 : X_0 \oplus X_1 \to Y_0 \oplus Y_1$ is Fredholm with $\ker(T_0 \oplus T_1) = (\ker T_0) \oplus (\ker T_1)$, $\text{ran}(T_0 \oplus T_1) = (\text{ran} T_0) \oplus (\text{ran} T_1)$ and so $\text{ind}(T_0 \oplus T_1) = \text{ind} T_0 + \text{ind} T_1$.

(3) If $K \in L(X)$ is compact and $c \neq 0$, then $K - cl$ is Fredholm and $\text{ind}(K - cl) = 0$ by the Riesz-Fredholm theorem and the lemma preceding it. (It is proved below that an operator is Fredholm with index 0 iff it is the sum of an invertible operator and a compact (in fact, finite rank) operator.)

(4) The unilateral shift on $l^2$ defined by $S(c_0, c_1, c_2, \ldots) = (0, c_0, c_1, c_2, \ldots)$ is Fredholm with $\text{ind} S = \dim \ker S - \text{codim} \text{ran} S = 0 - 1 = -1$. The backward shift $S^*$ on $l^2$ defined by $S^*(c_0, c_1, c_2, \ldots) = (c_1, c_2, c_3, \ldots)$ is also Fredholm with $\text{ind} S^* = \dim \ker S^* - \text{codim} \text{ran} S^* = 1 - 0 = 1$. (It is proved below that $\text{ind} T^* = -\text{ind} T$.) Also, $S^n$ and $(S^*)^n$ are Fredholm with $\text{ind}(S^n) = -n$ and $\text{ind}((S^*)^n) = n$. 46
(5) If \( T \in L(X,Y) \), \( \dim X < \infty \) and \( \dim Y < \infty \), then \( T \) is Fredholm with \( \text{codim ran } T = \dim(Y/\text{ran } T) = \dim Y - \dim \text{ran } T \), \( \text{dim ker } T + \dim \text{ran } T = \dim X \) from linear algebra and so \( \text{ind } T = \dim X - \dim Y \).

**Theorem (Atkinson).** Let \( T \in L(X,Y) \). The following are equivalent:

(a) \( T \) is Fredholm,

(b) there is \( S \in L(Y,X) \) such that \( I-TS \) and \( I-ST \) are finite rank (\( S \) is called a Fredholm inverse of \( T \)).

(c) there are \( S,S' \in L(Y,X) \) such that \( I-TS \) and \( I-ST \) are compact.

**Proof.** (a) \( \Rightarrow \) (b) Let \( P \in L(X) \) be a projection with \( \text{ran } P = \ker T \) and \( Q \in L(Y) \) be a projection with \( \text{ran } Q = \text{ran } T \). Let \( Z = \text{ran}(I-P) \). Since \( X = \text{ran } P \oplus \text{ran}(I-P) = \ker T \oplus Z \), \( T_0 = T|_Z : Z \to \text{ran } T \) is bijective. By the inverse mapping theorem, \( T_0 \) is invertible. Let \( S = T_0^{-1}Q : Y \to Z \subseteq X \). For all \( x \in X \), \( Tx \in \text{ran } Q \). So \( QTx = Tx = T(Px + (I-P)x) = T_0((I-P)x) \), i.e. \( QT = T_0(I-P) \). We have \( ST = T_0^{-1}QT = T_0^{-1}T_0(I-P) = I-P \) and \( TS = TT_0^{-1}Q = T_0T_0^{-1}Q = Q = I-(I-Q) \). Now \( \dim \text{ran } (I-ST) = \dim \text{ran } P = \dim \ker T < \infty \) and \( \dim \text{ran } (I-TS) = \dim \text{ran } (I-Q) = \text{codim } \text{ran } Q = \text{codim } \text{ran } T < \infty \).

(b) \( \Rightarrow \) (c) Let \( S' = S \). Finite rank operators are compact.

(c) \( \Rightarrow \) (a) \( TS = I + K \) for some compact operator \( K \in L(Y) \). By the lemma preceding the Riesz-Fredholm theorem, \( \text{ran } TS = \text{ran } (I+K) \) is closed and \( \text{codim } \text{ran } TS = \text{codim } \text{ran } (I+K) < \infty \). Also, since \( \text{ran } TS \subseteq \text{ran } T \subseteq Y \), \( \text{codim } \text{ran } T < \infty \). By the remarks following the definition of Fredholm operators, \( \text{ran } T \) is closed.

Next \( S'T = I + L \) for some compact operator \( L \in L(X) \). Since \( \dim \ker S'T = \dim \ker (I+L) < \infty \) and \( \ker T \subseteq \ker S'T \), we get \( \dim \ker T < \infty \). Therefore, \( T \) is Fredholm.

**Definition.** Let \( K(X) \) be the set of all compact operators on \( X \). By the properties of compact operators, we see \( K(X) \) is a closed two-sided ideal in \( L(X) \). Then \( L(X)/K(X) \) is a Banach space with a multiplicative structure and we called it the *Calkin algebra* on \( X \). (As in algebra, we define \( [T][S] = ([T]+K(X))(S+TK(X)) = TS+K(X) = [TS]\).

**Theorem (Simple Properties of Fredholm Operators).** (a) \( T \in L(X) \) is Fredholm iff \( [T] = T+K(X) \) is invertible in \( L(X)/K(X) \).

(b) If \( T \in L(X,Y) \) is Fredholm and \( K \in L(X,Y) \) is compact, then \( T+K \) is Fredholm.

(c) If \( T \in L(X,Y) \) is Fredholm and \( S \in L(Y,X) \) is a Fredholm inverse of \( T \), then \( S \) is Fredholm.

(d) If \( T \in L(X,Y) \) is Fredholm, then \( T^* \in L(Y^*,X^*) \) is Fredholm with \( \text{ind } T^* = -\text{ind } T \).

**Proof.** (a) If \( T \) is Fredholm, then let \( S \) be a Fredholm inverse of \( T \). We have \( [T][S]-[I] = [TS-I] = [0] = [ST-I] = [S][T]-[I] \). So \( [T][S] = [I] = [S][T] \). Conversely, if \( [S] = [T]^{-1} \in L(X)/K(X) \), then \( [I-STS] = [0] = [I-ST] \) implies \( I-STS \) and \( I-ST \) are compact. So \( T \) is Fredholm by Atkinson’s theorem.

(b) By (b) of Atkinson’s theorem, there is \( S \in L(Y,X) \) such that \( I-TS \) and \( I-ST \) are finite rank. Then \( I-(T+K)S = (I-TS)-KS \) and \( I-S(T+K) = (I-ST)-SK \) are compact, which implies \( T+K \) Fredholm by Atkinson’s theorem.

(c) This is clear from Atkinson’s theorem.

(d) By the closed range theorem, \( \text{ran } T \) closed implies \( \text{ran } T^* \) closed and \( \omega^* \)-closed. Since \( \ker T \) and \( Y/\text{ran } T \) are finite dimensional, by the kernel-range relations and the duality theorem,

\[
\dim \ker T^* = \dim(\text{ran } T)^\perp = \dim(Y/\text{ran } T)^* = \dim(Y/\text{ran } T) = \text{codim } \text{ran } T < \infty,
\]

\[
\text{codim } \text{ran } T^* = \text{codim } \text{ran } T^* = \text{codim } (\ker T)^\perp = \dim(X^*/(\ker T)^\perp) = \dim(\ker T)^* = \dim \ker T < \infty.
\]

Then \( \text{ind } T^* = \dim \ker T^* - \text{codim } \text{ran } T^* = \text{codim } \text{ran } T - \dim \ker T = -\text{ind } T \).
Lemma 1. If $T \in L(X,Y)$ is Fredholm and $M$ is a closed subspace of $X$, then $T(M)$ is closed in $Y$.  

Proof. As $\dim \ker T < \infty$, $\ker T$ has a complementary subspace $W$. Now $T|_W$ is injective and $\text{ran} \, T|_W = \text{ran} \, T$ is complete imply $T|_W$ is bounded below by the lower bound theorem. Thus, $T$ maps closed subspaces of $W$ to closed subspaces of $Y$. If $M$ is a closed subspace of $X$, then $T(M) = T(M + \ker T) = T((M + \ker T) \cap W)$ is a closed subspace of $Y$.

Lemma 2. If $F$ is a subspace of $X$ with finite codimension, $E_0$ is a subspace of $X$ such that $E_0 \cap F = \{0\}$, then there is a closed subspace $E \supseteq E_0$ such that $E \oplus F = X$.

Proof. For the quotient map $\pi : X \to X/F$, we have $\ker \pi = F$. So $\pi|_{E_0}$ is injective. Take a basis $B = \{x_1, \ldots, x_n\}$ of $E_0$. Then $\pi(B)$ is a basis of $\pi(E_0)$. Since $\dim (X/F) < \infty$, we can extend $\pi(B)$ to a basis $\{x_1 + F, \ldots, x_n + F\}$ of $X/F$ for some $n \geq i$. Let $E$ be the linear span of $\{x_1, \ldots, x_n\}$. Then $E$ contains $E_0$. Now $\dim E < \infty$ implies $E$ is complete, hence closed. Also, $\{x_1 + F, \ldots, x_n + F\}$ linearly independent implies $E \cap F = \{0\}$, while its span being $X/F$ implies $E + F = X$. Therefore, $E \oplus F = X$.

Multiplication Theorem. If $T \in L(X,Y)$ is Fredholm and $S \in L(Y,Z)$ is Fredholm, then $ST$ is Fredholm with $\text{ind}(ST) = \text{ind} S + \text{ind} T$.

Proof. (Due to Donald Sarason) In the case dim $X, \dim Y, \dim Z < \infty$, by example 5, $ST$ is Fredholm and $\text{ind}(ST) = \text{dim} X - \text{dim} Z = \text{dim} X - \text{dim} Y - \text{dim} Z = \text{ind} S + \text{ind} T$.

Otherwise, by lemma 1, $\text{ran} \, ST = S(\text{ran} \, T)$ is closed. Now dim ker $ST = \text{dim} T^{-1}(\text{ker} \, S) \leq \text{dim} \ker S + \text{dim} \ker T < \infty$ and codim ran $ST = \text{codim} S(\text{ran} \, T) \leq \text{codim} \, S + \text{codim} \, \text{ran} \, T < \infty$. So $ST$ is Fredholm.

(Our plan is to decompose $X = X_0 \oplus X_1, Y = Y_0 \oplus Y_1, Z = Z_0 \oplus Z_1$ with dim $X_0, \dim Y_0, \dim Z_0 < \infty$. Also, decompose $T = T|_{X_0} \oplus T|_{X_1}$, $S = S|_{Y_0} \oplus S|_{Y_1}$, where $T|_{X_1} : X_1 \to Y_1$ and $S|_{Y_1} : Y_1 \to Z_1$, with $T|_{X_1}$ and $S|_{Y_1}$ invertible. Then $ST|_{X_1} = S|_{Y_1} \circ T|_{X_1}: X_1 \to Z_1$ and $ST|_{X_1}$ is invertible. By examples 1 and 2, $\text{ind} S = \text{ind} S|_{Y_0}$, $\text{ind} T = \text{ind} T|_{X_0}$ and $\text{ind} (ST) = \text{ind}(ST|_{X_0})$.)

Let $X_0 = \ker ST$. From above, dim $X_0 < \infty$. So there is a closed subspace $X_1$ such that $X_0 \oplus X_1 = X$. By lemma 1, $Y_1 = TX_1$ is closed in $Y$. Since $\ker T \subseteq \ker ST = X_0$, so $\ker T \cap X_1 = \{0\}$ and $T|_{X_1} : X_1 \to TX_1 = Y_1$ is invertible. Now $\text{ran} T = TX_0 \oplus TX_1$ and $\text{dim} (\text{ran} T/\text{ran} X_1) = \text{dim} T X_0 \leq \text{dim} X_0 < \infty$. So

$$\text{codim} Y_1 = \text{dim}(Y/ TX_1) = \text{dim}(Y/ \text{ran} \, T) + \text{dim}(\text{ran} T/ TX_1) \leq \text{codim} \, \text{ran} \, T + \text{dim} X_0 < \infty.$$ (*)&

Next $\ker S \cap Y_1 = \ker S \cap TX_1 = \{0\}$ because $TX_1 \in \ker S$ for some $x_1 \in X_1$ implies $x_1 \in X_1 \cap X_0 = \{0\}$. By lemma 2, there is a closed subspace $Y_0 \supseteq \ker S$ such that $Y_0 \oplus Y_1 = Y$. Then $TX_0 = T(\text{ker} \, ST) = T(T^{-1}(\text{ker} \, S)) \subseteq \ker S \subseteq \text{Y}_0$, i.e. $T|_{X_0} : X_0 \to Y_0$. Also $\text{dim} Y_0 = \text{dim}(Y/Y_1) = \text{codim} Y_1 < \infty$. So $T = T|_{X_0} \oplus T|_{X_1}$.

By lemma 1, $Z_1 = SY_1$ is closed. Since $Y = Y_0 \oplus Y_1$ and $\ker S \subseteq Y_0$, so $\ker S \cap Y_1 = \{0\}$ and $S|_{Y_1} : Y_1 \to SY_1 = Z_1$ is invertible. As in (*), above, codim $Z_1 \leq \text{codim} \, \text{ran} \, S + \dim Y_0 < \infty$.

Next $SY_0 \cap Z_1 = SY_0 \cap SY_1 = \{0\}$ because $Sy_0 = Sy_1$ for $y_0 \in Y_0, y_1 \in Y_1$ implies $y_0 - y_1 \in \ker S \subseteq \text{Y}_0$, which implies $y_1 \in \text{Y}_0 \cap Y_1 = \{0\}$, then $Sy_0 = Sy_1 = 0$. By lemma 2, there is a closed subspace $Z_0 \supseteq SY_0$ such that $Z_0 \cap Z_1 = Z$. Also, dim $Z_0 = \text{codim} Z_1 < \infty$. So $S = S|_{Y_0} \oplus S|_{Y_2}$.

By examples 1 and 2, $\text{ind} S = \text{ind} S|_{Y_0}$, $\text{ind} T = \text{ind} T|_{X_0}$ and $\text{ind} (ST) = \text{ind}(ST|_{X_0})$. Since $ST|_{X_0} = S|_{Y_0} T|_{X_0}$ and dim $X_0, \dim Y_0, \dim Z_0 < \infty$, the theorem now follows from the finite dimensional case.

Perturbation Theorem. Let $T \in L(X,Y)$ be Fredholm. Then there is $\varepsilon > 0$ so that $T + A$ is Fredholm with $\text{ind}(T + A) = \text{ind} T$, where $A \in L(X,Y)$ and $\|A\| < \varepsilon$. (This implies the Fredholm operators form an open set in $L(X,Y)$ and the index is continuous and constant on each connected component of that set.)

Proof. Let $S \in L(Y,X)$ be such that $K = I - TS$ and $L = I - ST$ are finite rank. Let $\varepsilon = \|S\|^{-1}$. Assume $\|A\| < \varepsilon$. Since $\|AS\| \leq \|A\| \|S\| < 1$, $I + AS$ is invertible. Now

$$(T + A)S = I - K + AS = (I + AS) - K = \left( I - K(I + AS)^{-1} \right) (I + AS).$$

48
Solving for \( K(I + AS)^{-1} \), we see

\[
I - (T + A) \left( S(I + AS)^{-1} \right) = K(I + AS)^{-1}
\]

is compact. Similarly, \( I + SA \) is invertible and \( I - \left( (I + SA)^{-1} S \right)(T + A) = (I + SA)^{-1}L \) is compact. So, by Atkinson’s theorem, \( T + A \) is Fredholm. By example 3 and the multiplication theorem,

\[
0 = \text{ind}(I - K(I + AS)^{-1}) = \text{ind}(T + A)S(I + AS)^{-1} = \text{ind}(T + A) + \text{ind}S + \text{ind}(I + AS)^{-1} = \text{ind}(T + A) + \text{ind}S.
\]

Also, \( 0 = \text{ind}(I - K) = \text{ind}(TS) = \text{ind}T + \text{ind}S \). Therefore, \( \text{ind}(T + A) = \text{ind}T \).

**Corollary.** If \( T \in L(X, Y) \) is Fredholm and \( K \in L(X, Y) \) is compact, then \( \text{ind}(T + K) = \text{ind}T \).

**Proof.** Since \( f(t) = \text{ind}(T + tK) \) is a continuous function on \([0, 1]\) with integer value, it is a constant function. In particular, \( \text{ind}(T + K) = f(1) = f(0) = \text{ind}(T) \). \(\square\)

**Theorem.** Let \( A \in L(X, Y) \). The following are equivalent.

(a) \( A \) is Fredholm with \( \text{ind}A = 0 \),

(b) \( A = C + F \), where \( C \) is invertible in \( L(X, Y) \) and \( F \) is finite rank in \( L(X, Y) \),

(c) \( A = B + K \), where \( B \) is invertible in \( L(X, Y) \) and \( K \) is compact in \( L(X, Y) \).

**Proof.** (a) \(\Rightarrow\) (b) If \( \text{ind}A = 0 \), then \( \dim \ker A = \text{codim ran} A < \infty \). Let \( Z \) be a complementary subspace of \( \ker A \) in \( X \). Let \( W \) be a complementary subspace of \( \text{ran} A \) in \( Y \). Let \( P \in L(X) \) be a projection such that \( \text{ran} P = \ker A \) is finite dimensional. Since \( \dim W = \text{dim} \text{ran} A = \dim \ker A < \infty \), there is an invertible operator \( T : \ker A \to W \).

Now \( A + TP \) is injective because \( (A + TP)(x) = 0 \) implies \( Ax = -TPx \in \text{ran} A \cap W = \{0\} \). Then \( Ax = 0 \) implies \( x \in \ker A = \text{ran} P \) so that \( Px = x \) and \( TX = T(=0) = 0 \). Since \( T \) is invertible, \( x = 0 \).

For surjectivity of \( A + TP \), first observe that \( X = \ker A \oplus Z \) implies \( \text{ran} A = A(X) = A(Z) \). Next, \( P \) is the projection onto \( \ker A \) implies \( P(Z) = \{0\} \). Also, \( TP(\ker A) = TP(\text{ran} P) = T(\text{ran} P) = T(\ker A) = W \). Then, \( A + TP \) is surjective since \( (A + TP)(\ker A \oplus Z) = TP(\ker A) \oplus A(Z) = W \oplus \text{ran} A = Y \).

Hence, \( A + TP \) is invertible. Since \( \dim W < \infty \), \( TP \) is finite rank. Then \( A = (A + TP) - TP \) satisfies the required conditions.

(b) \(\Rightarrow\) (c) Finite rank implies compactness.

(c) \(\Rightarrow\) (a) \( B + K \) is Fredholm follows by example 1 and part (b) of the simple properties of Fredholm operators. Also, by example 1 and the last corollary, \( \text{ind}(B + K) = \text{ind}(B) = 0 \). Alternatively, \( \text{ind}A = \text{ind}(B + K) = \text{ind}B(I + B^{-1}K) = \text{ind}B + \text{ind}(I + B^{-1}K) = 0 \) by the multiplication theorem, examples 1 and 3. \(\square\)

Throughout this chapter $H, H_1, H_2$ will denote Hilbert spaces over $\mathbb{C}$. The inner product on $H$ will be denoted by $\langle \cdot, \cdot \rangle$. For every $y \in H$, the linear functional $f_y(x) = \langle x, y \rangle$ is in $H^*$. Recall that the Riesz representation theorem asserted that there is a bijection from $H$ onto $H^*$ given by $y \mapsto f_y$. For all $y, y' \in H$, it satisfies $\|y\| = \|f_y\|$, $f_y + f_{y'} = f_{y+y'}$. It may seem $H^*$ is isometric isomorphic to $H$. Unfortunately, for all $c \in \mathbb{K}$ and $y \in H$, $f_{cy} = c f_y$. Keeping this in mind, we say there is a conjugate-linear isometric isomorphism between $H$ and $H^*$. By a slight abuse of meaning, it is popular to write $H^* = H$, where $f_y$ is identified with $y$. In particular, $H$ is reflexive so that the weak and weak-star topologies coincide.

Now for every $T \in \text{L}(H_1, H_2)$ and $y \in H_2$, the function $f(x) = \langle Tx, y \rangle$ is in $H_1^*$.

Remark. Throughout this chapter $H, H_1, H_2$ be given such that for every $x \in H_1$ such that $f(x) = \langle x, w \rangle$. Define the adjoint of $T \in \text{L}(H_1, H_2)$ to be $T^* \in \text{L}(H_2, H_1)$ given by $T^* y = w$. So $(Tx, y) = \langle x, T^*y \rangle$ for all $x \in H_1, y \in H_2$. In particular, for $T \in \text{L}(H_1, H_2)$, $T^{**} = T$. This looks very similar to the adjoint for Banach spaces. Again, there is one exception, namely $(cT)^* = \overline{c} T^*$.

In general, facts about Banach spaces also apply to Hilbert spaces and in some places where adjoints were needed, we need to do conjugations. For example, $(T - cI)^* = T^* - \overline{c} I$. So $\sigma(T^*) = \{ \overline{c} : c \in \sigma(T) \}$.

Definitions. (1) An involution on a Banach algebra $B$ is a map from $B$ to $B$ sending every $x \in B$ to some $x^* \in B$ such that for every $a, b \in B$ and $c \in \mathbb{K}$, $a x^* = \overline{a} (a^* x^*) = (a + b)^* = a^* + b^*$ and $(c a)^* = \overline{c} a^*$.

(2) A $C^*$-algebra is a Banach algebra $B$ with an involution such that for every $x \in B$, we have $\|x^* x\| = \|x\|^2$. (Note $\|x^*\| = \|x\|$ because $\|x^* x\| \leq \|x^*\| \|x\| \|x\|$ implies $\|x\| \leq \|x^*\|$ and from this, $\|x^*\| \leq \|x^* x\| = \|x\|$.

Theorem. For $T \in \text{L}(H_1, H_2)$, we have $\|T^* T\| = \|T\|^2$. (So $\text{L}(H)$ is a $C^*$-algebra with adjoint as involution.) Also, $H_1 = \ker T \oplus \text{ran} T^*$ and $H_2 = \ker T^* \oplus \text{ran} T$.

Proof. Since $\|T^*\| = \|T\|$, so $\|T^* T\| \leq \|T^*\| \|T\| = \|T\|^2$. Conversely, for $\|x\| \leq 1$, $\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^* T x, x \rangle \leq \|T^* T\| \|x\|^2 \leq \|T^* T\| \|T\|^2 \|x\|^2$, which implies $\|T\|^2 \leq \|T^* T\|$. The last statement follows from $H = V \oplus V^\perp$ in a Hilbert space $H$ with a subspace $V$ and the formulas $(\ker T)^\perp = \text{ran} T^*$ and $(\ker T^*)^\perp = \text{ran} T$.

Recall the projection theorem asserts that for every subspace $M$ of $H$, every $x \in H$ has a unique decomposition $x = y + z$, where $y \in M$ (is the closest point to $x$ in $M$) and $z \in M^\perp$. The function $P_M : H \to M$ defined by $P_M(x) = y$ is a projection since $P_M^2 x = P_M y = y = P_M x$. Its kernel $M^\perp$ and its range $M$ are orthogonal. If $M \neq \{0\}$, then $\|P_M\| = 1$. Note $P_M^M = I - P_M$ and ker $P_M = M = \text{ran} P_M = \text{ran}(I - P_M)$.

Definition. A projection $P \in \text{L}(H)$ is orthogonal iff $\ker P \perp \text{ran} P$. In that case, $P = P_M$, where $M = \text{ran} P$.

Theorem. For a nonzero projection $P$, (a) $P$ is orthogonal, (b) $P^* = P$ and (c) $\|P\| = 1$ are equivalent.

Proof. (a) $\Rightarrow$ (b) $P$ is orthogonal implies $\text{ran} P \perp \text{ran}(I - P)$. So, for all $x \in H$, $0 = \langle Px, (I - P)x \rangle = \langle (I - P^*) P x, x \rangle$. Then $(I - P^*) P = 0$, i.e. $P = P^* P$. So $P^* = (P^* P)^* = P^* P^* = P^* P = P$.

(b) $\Rightarrow$ (c) $P^* = P$ implies $\|P x\|^2 = \langle P x, P x \rangle = \langle P^* P x, x \rangle = \langle P^2 x, x \rangle = \langle P x, x \rangle \leq \|P x\| \|x\|$. So $\|P x\| \leq \|x\|$ with equality if $x \in \text{ran} P$. Thus, $\|P\| = 1$.

(c) $\Rightarrow$ (a) Assume $P$ is not orthogonal. Then there is $x \in \text{ran} P$, $y \in \ker P$ such that $\|x\| = 1 = \|y\|$ and $\langle x, y \rangle \neq 0$. Replacing $x$ by $e^{i\theta} x$, we may assume $\langle x, y \rangle = -t < 0$. Take $z = x + ty$ Then $\|z\|^2 = \|x\|^2 + 2t(x, y) + t^2 \|y\|^2 = 1 - t^2 < 1 = \|x\|^2 = \|Pz\|^2$, which implies $\|P\| \neq 1$, contradiction.

Remark. For an orthogonal projection $P$, in the last proof we saw $\|P x\|^2 = \|P x\|^2$. This is useful.

Theorem (Sum of Orthogonal Projections). Let $E, F$ be orthogonal projections with ranges $Y, Z$, respectively. The following are equivalent:

50
(a) \( Y \perp Z \), (b) \( E(Z) = \{0\} \), (c) \( EF = 0 \), (d) \( F(Y) = \{0\} \) and (e) \( FE = 0 \).

Also \( E + F \) is an orthogonal projection iff \( Y \perp Z \), in which case \( \text{ran}(E + F) = Y + Z \) is the closed linear span of \( Y \cup Z \).

**Proof.** \( Y \perp Z \Leftrightarrow Z \subseteq Y^\perp = \ker E \Leftrightarrow E(\text{ran} F) = \{0\} \Leftrightarrow E(Fx) = 0 \) for all \( x \in H \Leftrightarrow EF = 0 \). Similarly \( Z \perp Y \Leftrightarrow F(Y) = \{0\} \Leftrightarrow FE = 0 \).

If \( Y \perp Z \), then \( (E + F)^2 = E^2 + EF + FE + F^2 = E + 0 + 0 + F = E + F \) and \( (E + F)^* = E^* + F^* = E + F \), so \( E + F \) is an orthogonal projection.

Conversely, \( E + F \) is an orthogonal projection implies \( \|E + F\| = 1 \). So for \( x \in Y \),

\[
\|x\|^2 \geq \|(E + F)x\|^2 = \|(E,F)x\| = \|Ex\|^2 + \|Fx\|^2 = \|x\|^2 + \|Fx\|^2.
\]

So \( F(Y) = \{0\} \), which is equivalent to \( Y \perp Z \).

Finally, in case \( E + F \) is an orthogonal projection, let \( M \) be the closed linear span of \( Y \cup Z \). Since \( (E + F)|_Y = E|_Y + 0 = I \) and similarly \( (E + F)|_Z = I \), we have \( (E + F)|_{Y \cup Z} = I \). Then \( M \subseteq \text{ran}(E + F) \subseteq Y + Z \subseteq M \). So \( \text{ran}(E + F) = Y + Z = M \). \( \square \)

**Exercises.** Let \( E, F \) be orthogonal projections with ranges \( Y, Z \), respectively.

(1) Prove that \( E \) and \( F \) are orthogonal projections if \( EF = FE \), in which case, \( \text{ran}(E + F) = Y \cap Z \).

(2) Prove that the following are equivalent: (a) \( Y \subseteq Z \), (b) \( FE = E \), (c) \( EF = E \), (d) \( \|Ex\| \leq \|Fx\| \) for all \( x \in H \) and (e) \( E \leq F \). Then prove that \( F - E \) is an orthogonal projection iff \( Y \subseteq Z \), in which case \( \text{ran}(F - E) = Z \cap Y^\perp \).

**Definitions.** The **numerical range** of \( T \in L(H) \) is \( V(T) = \{\langle Tx, x \rangle : \|x\| = 1\} \). The **numerical radius** of \( T \) is \( \sup\{\|Tx, x\| : \|x\| = 1\} \).

**Theorem.** Let \( T, T_0, T_1 \in L(H) \).

1. \( T = 0 \) if and only if \( V(T) = \{0\} \), i.e., \( \langle Tx, x \rangle = 0 \) for all \( x \in H \). \( T_0 = T_1 \) if and only if \( \langle T_0x, x \rangle = \langle T_1x, x \rangle \) for all \( x \in H \).

2. \( \sigma(T) \subseteq V(T) \) and if the distance from \( c \) to \( V(T) \) is \( d > 0 \), then \( \|T - cI\| \leq 1/d \).

**Proof.** (1) \( T = 0 \) implies \( V(T) = \{0\} \) is trivial. For the converse, let \( x, y \in H \), then

\[
\langle Tx, y \rangle = \frac{1}{4} \left( \langle (T(x + y), x + y) \rangle - \langle (T(x - y), x - y) \rangle + i\langle (T(x + iy), x + iy) \rangle - i\langle (T(x - iy), x - iy) \rangle \right) = 0.
\]

In the case \( y = Tx \), we get \( \|Tx\|^2 = \langle Tx, Tx \rangle = 0 \). Then \( T = 0 \).

(2) Let \( c \notin V(T) \). Then the distance from \( c \) to \( V(T) \) is \( d > 0 \). For \( \|x\| = 1 \), \( \|(T - cI)x\| \geq \|\langle (T - cI)x, x\rangle\| = \|\langle Tx, x\rangle - c\| \geq d > 0 \) implies \( T - cI \) is bounded below. By the lower bound theorem, \( T - cI \) is injective and has closed range. Assume \( \text{ran}(T - cI) \) is not dense. Then \( \ker(T^* - cI) = (\text{ran}(T - cI))^\perp \neq \{0\} \). So there is \( \|v\| = 1 \) such that \( T^*v = cv \). Then \( c = (v, cv) = (v, T^*v) = (Tv, v) \in V(T) \), a contradiction. Hence \( \text{ran}(T - cI) \) is dense. So \( T - cI \) is invertible and \( c \notin \sigma(T) \). From the inequality for \( T - cI \) bounded below, we get \( \|T - cI\| \leq \|c - c\| \leq 1/d \).

**Remark.** The Toeplitz-Hausdorff theorem asserts that \( V(T) \) is convex. See [BN], pp. 387-389.

**Definitions.** Let \( T \in L(H) \).

(1) \( T \) is **normal** if \( T^*T = TT^* \). (This is equivalent to \( \|Tx\| = \|T^*x\| \) for all \( x \in H \) because \( \|(T^*T - TT^*)x, x\rangle = \|Tx\|^2 - \|T^*x\|^2 \). So \( \ker T = \ker T^* \) and \( \text{ran} T = (\ker T^*)^\perp = (\ker T)^\perp = \text{ran} T^* \).

(2) \( T \) is **self-adjoint** (or **Hermitian**) if \( T^* = T \). (This is equivalent to \( \langle Tx, x \rangle \in \mathbb{R} \) for all \( x \in H \) because \( \|(T - T^*)x, x\rangle = \langle (Tx, x) - (T^*x, x) \rangle = \langle (Tx, x) - (x, Tx) \rangle = 2i \text{ Im}(\langle Tx, x \rangle). \)
(3) $T$ is positive (and we write $T \geq 0$) iff $(Tx,x) \geq 0$ for all $x \in X$. (In this case, $(Tx,x) = (x,Tx) = (T^*x,x)$ implies $T^* = T$.) For self-adjoint operators $A$ and $B$, define $A \leq B$ (or $B \geq A$) iff $B - A \geq 0$.

(4) $T$ is an isometry iff $T^*T = I$. (This is equivalent to $\|Tx\| = \|x\|$ for all $x \in X$ because $(T^*T - I)x,x = (T^*T,x) - (x,x) = \|Tx\|^2 - \|x\|^2$.)

(5) $T$ is unitary iff $TT^* = I = T^*T$. (By (4), it is equivalent to an invertible isometry.)

Other than isometry, these are all normal operators. Also, for orthogonal projection $P$, since $(Px,x) = \|Px\|^2 \geq 0$, they are positive, hence normal. Now we begin to study normal operators.

**Theorem (Basic Properties of Normal Operators).** Let $T \in L(H)$ be normal.

(1) For every $c \in \mathbb{C}$, $T - cl$ is normal.

(2) Eigenvectors for different eigenvalues of $T$ are orthogonal, i.e. if $a \neq b$, $Tx = ax$ and $Ty = by$, then $(x,y) = 0$.

(3) $T$ is invertible iff $T$ is right invertible iff $T$ is bounded below iff $T$ is left invertible.

(4) $\sigma(T) = \sigma_{ap}(T)$.

(5) The spectral radius and the numerical radius both equal $\|T\|$.  

**Proof.** (1) $(T - cl)(T - cl)^* = (T - cl)(T^* - \overline{c}I) = TT^* - cT^* - \overline{c}T + |c|^2 = T^*T - cT^* - \overline{c}T + |c|^2 = (T^* - \overline{c}I)(T - cl) = (T - cl)^*(T - cl)$.

(2) Since $0 = \|T - bI\|y = \|(T - bI)y\| = \|(T^* - \overline{b}I)y\|$, so $T^*y = \overline{b}y$. Then $(x,y) = (Tx,y) = (x,T^*y) = (x,\overline{b}y) = b(x,y)$ and $a \neq b$ imply $(x,y) = 0$.

(3) Note $T$ is right invertible $\iff T^*$ is left invertible $\iff T^*$ is bounded below $\iff T$ is bounded below $\iff T$ is left invertible. Finally $T$ invertible $\iff T$ is right invertible $\iff T$ is left and right invertible $\iff T$ is invertible.

(4) By (1) and (3), $c \notin \sigma(T)$ iff $T - cl$ is invertible iff $T - cl$ is bounded below iff $c \notin \sigma_{ap}(T)$.

(5) $\|T^2\| = \|(T^2)^*T^2\|^{1/2} = \|(T^*T)(T^*T)\|^{1/2} = \|T^*T\| = \|T\|^2$. Iterating this, we get $\|T^{2n}\| = \|T\|^{2n}$. Therefore, $r(T) = \lim_{n \to \infty} \|T^{2n}\|^{1/2n} = \|T\|$. 

Next, since $\sigma(T)$ is compact, there is $c \in \sigma(T)$ with $|c| = r(T) = \|T\|$. By (4), there are $x_n \in X$ such that $\|x_n\| = 1$ and $\|(T - cl)x_n\| \to 0$. Since $\|(T - cl)x_n\| \geq \|((T - cl)x_n,x_n)\| = \|Tx_n,x_n\| - |c|$, so $(Tx_n,x_n) \to c$. Hence $\|T\| = |c| = \lim\sup_{n \to 0} \|(Tx_n,x_n)\| \leq \sup\{(Tx,x) : \|x\| = 1\} \leq \|T\|$ and the numerical radius of $T$ is $\|T\|$.  

**Remark.** It is known that for a normal operator, the closure of the numerical range is the convex hull of the spectrum. See [H], pp. 116 and 318.

**Theorem.** (1) If $T$ is self-adjoint, then $\sigma(T) \subseteq \mathbb{R}$ and for $c \notin \mathbb{R}$, $\|(T - cl)^{-1}\| \leq 1/|\text{Im} c|$.

(2) If $T$ is unitary, then $\sigma(T) \subseteq \{z : |z| = 1\}$ and for $c \neq 0$, $\|(T - cl)^{-1}\| \leq 1/|1 - |c||$.

(3) If $T \geq 0$, then $\sigma(T) \subseteq [0, +\infty)$ and for $c \notin \mathbb{R}$, $\|(T - cl)^{-1}\| \leq 1/\text{Im} c$ and for $c < 0$, $\|(T - cl)^{-1}\| \leq 1/|c|$.

**Proof.** (1) Since $(Tx,x) \in \mathbb{R}$, we get $\sigma(T) \subseteq \overline{\mathbb{V}(T)} \subseteq \mathbb{R}$. If $c \notin \mathbb{R}$, then $T - cl$ is invertible and we have $\|(T - cl)x\| \geq \|(T - cl)x,x\| \geq \|T^*x,x\| \geq \text{Im} c \|x\|^2$. Setting $y = (T - cl)x$, we get $\|y\| \geq \|\text{Im} c\|(T - cl)^{-1}y\|$, which implies $\|(T - cl)^{-1}\| \leq 1/\text{Im} c$.

(2) For $c \neq 0$, since $\|Tx\| = \|x\|$, we have $\|(T - cl)x\| \geq \|Tx - cx\| = |1 - |c|| \|x\|$. So $T - cl$ is normal and bounded below, hence invertible. We get $\sigma(T) \subseteq \{z : |z| = 1\}$. The norm estimate of $(T - cl)^{-1}$ is similar to that in (1).

(3) Since $(Tx,x) \geq 0$, we get $\sigma(T) \subseteq \overline{\mathbb{V}(T)} \subseteq [0, +\infty)$. For $c \notin \mathbb{R}$, we can repeat the reasoning in (1). For $c < 0$, we have $\|(T - cl)x\| \geq \|Tx - cx\| \geq |1 - |c|| \|x\|$, $\|c\| \|x\|^2$ and the rest is similar.  

**Exercise.** If $T$ is normal and $c \notin \sigma(T)$, then prove that $\|(T - cl)^{-1}\| = 1/\text{inf}\{|z - c : z \in \sigma(T)\}$.
Theorem. If $T$ is self-adjoint, then
(1) either $\|T\|$ or $-\|T\|$ is in $\sigma(T)$,
(2) $\sup \sigma(T) = \sup V(T),$ inf $\sigma(T) = \inf V(T)$ and so $\sigma(T) \subseteq [\inf \sigma(T), \sup \sigma(T)] = [\inf V(T), \sup V(T)]$ (in particular, $m = \inf V(T)$ and $M = \sup V(T)$ are in $\sigma(T) = \sigma_{ap}(T)$),
(3) $T \geq 0$ iff $\sigma(T) \subseteq [0, +\infty)$.

Proof. (1) By property (5) of normal operators, $r(T) = \|T\|$. Since $\sigma(T) \subseteq \mathbb{R}$ and $\{z \in \mathbb{C} : |z| = r(T)\}$ intersects $\sigma(T)$, so either $\|T\|$ or $-\|T\|$ is in $\sigma(T)$.

(2) Let $M = \sup V(T) = \sup \{(Tx, x) : \|x\| = 1\}$ and $M' = \sup \sigma(T) = \sup \{c : c \in \sigma(T)\}$. Now $S = |T|I + T$ is positive (as $\langle(T|I + T)x, x\rangle = \|T\|^2 \|x\|^2 + \langle Tx, x\rangle \geq 0$) and self-adjoint. Now $V(S) = \{\|T\| + \langle Tx, x\rangle : \|x\| = 1\} \subseteq [0, +\infty)$ and $\sigma(S) = \{(|T| + c : c \in \sigma(T))\} \subseteq [0, +\infty)$. By property (5) of normal operators, the numerical radius of $S$ and the spectral radius of $S$ are equal. So $\|T\| + M = \|T\| + M'$. Hence $M = M'$. Applying a similar argument to $\|T\|I - T$, we see the infima are the same.

(3) The only-if direction follows from part (3) of the last theorem. For the if-direction, since $\sigma(T) \subseteq [0, +\infty)$, so by (2), $\inf V(T) = \inf \sigma(T) \geq 0$. Then $V(T) \subseteq [\inf V(T), \infty) \subseteq [0, +\infty)$, which implies $T \geq 0$.

Definitions. Let $T \in \mathcal{L}(H)$ and $M$ be a subspace of $H$. We say $M$ is invariant under $T$ iff $T(M) \subseteq M$. Also, $M$ reduces $T$ iff $T(M) \subseteq M$ and $T(M^\perp) \subseteq M^\perp$.

Lemma. Let $T \in \mathcal{L}(H), M$ be a subspace of $H$ and $P$ be the orthogonal projection onto $M$.
(1) $M$ is invariant under $T$ iff $PTP = TP$ iff $M^\perp$ is invariant under $T^*$.
(2) $M$ reduces $T$ iff $PTP = TP$ iff $T(M) \subseteq M$ and $T^*(M) \subseteq M$ iff $M$ reduces $T^*$.

Proof. (1) For $x \in H$, write $x = y + y'$, $Ty = z + z'$, where $y, z \in M$, $y', z' \in M^\perp$. We have $PTPx = PTy = z$ and $TPx = Ty = z + z'$. So $PTP = TP$ iff $Ty \in M$ for all $x \in H$ iff $T(M) \subseteq M$.

Next $Q = I - P$ is the orthogonal projection onto $M^\perp$. So $T^*(M^\perp) \subseteq M^\perp$ iff $(I - P)T^*(I - P) = T^*(I - P)$, which expands and simplifies to $PT^*P = PT^*$. Finally, by taking adjoint of both sides, $PT^*P = PT^*$ is equivalent to $PTP = TP$.

By (1), since $I - P$ is the orthogonal projection onto $M^\perp$, $M$ reduces $T$ iff $PTP = TP$ and $(I - P)T(I - P) = T(I - P)$ iff $PTP = TP$ and $PT^*P = PT^*P$ iff $PT^*P = T^*P$ iff $M$ reduces $T^*$.

Remarks. For all $x, y \in M$, $(x, (T|M)^*y) = ⟨T|Mx, y⟩ = ⟨Tx, y⟩ = ⟨x, T^*y⟩ = ⟨x, T^*|M)y⟩$. So $(T|M)^* = T^*|M$. Similarly, $(T|M^\perp)^* = T^*|M^\perp$.

Theorem (Further Properties of Normal Operators). Let $T \in \mathcal{L}(H)$ be normal.
(6) For every $c \in \mathbb{C}$, ker($T - cI$) reduces $T$ (and hence also $T^*$).

(7) For a subspace $M$ of $H$, if $M$ reduces $T$, then $T|M$, $T|M^\perp$ and their adjoints are normal and $\|T\| = \max\{\|T|M\|, \|T|M^\perp\}$.

Proof. (6) For $x \in \ker(T - cI)$, $(T - cI)Tx = T(T - cI)x = 0$ implies $Tx \in \ker(T - cI)$. Similarly, $(T - cI)T^*x = T^*(T - cI)x = 0$ implies $T^*x \in \ker(T - cI)$. By the lemma, ker($T - cI$) reduces $T$ and $T^*$.

(7) Using the remark, $T|M(T|M)^* = T|M(T^*M) = (TT^*)|M = (T^*T)|M = T^*|MT|M = (T|M)^*T|M$. So $T|M$ and $T^*|M$ are normal. Since $M^\perp$ also reduces $T$, similarly $T|M^\perp$ and $T^*|M^\perp$ are normal.

Next, clearly $\|T|M\| = \|T|M^\perp\| \leq \|T\|$. So max$\{\|T|M\|, \|T|M^\perp\}$ $\leq \|T\|$. For the reverse inequality, write $x = y + z$, where $y \in M$ and $z \in M^\perp$. Then $\|x\|^2 = \|y\|^2 + \|z\|^2$. Since $M$ reduces $T$, so $Ty \in M, Tz \in M^\perp$. Then $\|Tx\|^2 = \|Ty\|^2 + \|Tz\|^2 \leq \|T|M\|^2\|y\|^2 + \|T|M^\perp\|^2\|z\|^2 \leq \max\{\|T|M\|, \|T|M^\perp\}\|x\|^2$.

So $\|T\| \leq \max\{\|T|M\|, \|T|M^\perp\}$.
**Spectral Theorem for Compact Normal Operators.** Let \( T \in L(H) \) be a compact normal operator. For an eigenvalue \( c \) of \( T \), let \( P_c \) denote the orthogonal projection onto \( H_c = \ker(T - cI) \). As \( \sigma(T) \) is a countable set with \( 0 \) as the only possible accumulation point, let its nonzero elements be \( c_1, c_2, c_3, \ldots \) arranged so that \(|c_1| \geq |c_2| \geq |c_3| \geq \cdots\). Then \( T = \sum c_i P_{c_i} \) (where the series converges in the norm of \( L(H) \) if there are infinitely many terms) and \( H \) has an orthonormal basis consisting of eigenvectors of \( T \).

**Proof.** Since \( T \) is compact, the \( H_c \)'s (\( c \neq 0 \)) are finite dimensional. Since \( T \) is normal, by property (2) of normal operators, the \( H_c \)'s (for all \( c \in \sigma(T) \)) are pairwise orthogonal, i.e. \( P_c P_{c'} = 0 \) if \( c \neq c' \).

For every \( \varepsilon > 0 \), there is \( N \) such that \( \sigma(T) \setminus \{c_1, c_2, \ldots, c_N\} \subseteq B(0, \varepsilon) \). Let \( M = \sum_{i=1}^N H_{c_i} \) and \( T_N = \sum_{i=1}^N c_i P_{c_i} \).

By property (6) of normal operators, \( H_c \) reduces \( T \) (i.e. \( T P_c = P_c T \)). Also \( T_N P_c = P_c T_N \) (i.e. \( H_c \) reduces \( T_N \)). Since \( H_c \)'s are pairwise orthogonal, \( P_M = \sum_{i=1}^N P_{c_i} \) by the theorem on sum of orthogonal projections.

Then \( M \) reduces \( T \) and \( T_N \). Note \( T|_M = T_N|_M \) (as \( T|_M (v_i) = c_i v_i = T_N|_M (v_i) \) for all \( v_i \in H_{c_i} \)) and \( T_N|_{M^\perp} = 0 \) (as \( v \in M^\perp \) implies \( v \perp H_{c_i} \) and so \( P_{c_i} (v) = 0 \)). By property (7) of normal operators,

\[
\|T - T_N\| = \max\{\|T|_M - T_N|_M\|, \|T|_{M^\perp} - T_N|_{M^\perp}\} = \|T|_{M^\perp}\|.
\]

By property (7) of normal operators and properties (1) of compact operators, \( T|_{M^\perp} \) is also a compact normal operator. By the definition of \( M \), the eigenvalues of \( T|_{M^\perp} \) are in \( \sigma(T) \setminus \{c_1, c_2, \ldots, c_N\} \) and so \( \|T|_{M^\perp}\| = r(T|_{M^\perp}) < \varepsilon \). Therefore, \( T \) is the limit of \( T_N \) in the norm of \( L(H) \) and so \( T = \sum c_i P_{c_i} \).

Let \( H' \) be the closed linear span of all \( H_c \)'s, where \( c \in \sigma(T) \). Since \( H_c \)'s reduce \( T \), we have \( T^* (H_c) \subseteq H_c \) and so \( T^* (H') \subseteq H' \) and \( T (H'\perp) \subseteq H'^\perp \). Then \( T|_{H'^\perp} \) is compact normal and cannot have any nonzero eigenvalues by the definition of \( H' \). So \( \sigma(T|_{H'^\perp}) = \{0\} \) and \( \|T|_{H'^\perp}\| = r(T|_{H'^\perp}) = 0 \). Then \( H'^\perp \subseteq \ker T = H_0 \). By the definition of \( H' \), \( H'^\perp \cap H_0 = \{0\} \). So \( H'^\perp = \{0\} \). Therefore \( H' = H \) and taking an orthonormal basis in every \( H_c \) (\( c \in \sigma(T) \)), their union forms an orthonormal basis of \( H' = H \).

**Proof.** Apply the spectral theorem to \( T_1 \). Then \( H \) is the closed linear span of all \( H_c = \ker(T_1 - cI) \), where \( c \in \sigma(T_1) \). Since \( x \in H_c \) implies \( (T_1 - cI)T_2 x = T_2 (T_1 - cI) x = 0 \) (i.e. \( T_2 x \in H_c \)), it follows that each \( H_c \) is invariant under \( T_2 \). Then apply the spectral theorem to \( T_2 \) on every \( H_c \). The union of the orthonormal bases of \( H_c \) is a desired orthonormal basis for \( H \).

**Tensor Notations for Rank One Operators.** For \( v, e \in H \), define the linear functional \( e \otimes v \) on \( H \) by \((e \otimes v)(x) = (x, v) e \). If \( v, e \neq 0 \), then it is a rank one operator since its range is the span of \( \{e\} \).

**Theorem.** Every rank \( n \) operator \( F \in L(H) \) is the sum of \( n \) rank one operators.

**Proof.** Let \( \{e_1, \ldots, e_n\} \) be an orthonormal basis of \( \text{ran} \ F \). Since \( e_i \in \text{ran} \ F \), \( g_i(x) = (F(x), e_i) \) is a nonzero element of \( H^* \). By the Riesz representation theorem, there is a nonzero \( v_i \in H \) such that \( g_i(x) = (x, v_i) \). Then \( F(x) = \sum_{i=1}^n F(x, e_i) e_i = \sum_{i=1}^n g_i(x) e_i = \sum_{i=1}^n (x, v_i) e_i \), i.e. \( F = \sum_{i=1}^n e_i \otimes v_i \).

**Theorem.** Let \( T \in L(H) \) be a compact operator. Then there are countable orthonormal sets \( \{e_i\} \) and \( \{v_i\} \) in \( H \) and positive real numbers \( \{c_i\} \) converging to \( 0 \) if infinitely many such that for all \( x \in H \),

\[
T x = \sum_i c_i (x, v_i) e_i.
\]

The \( c_i \)'s are called the singular values of \( T \). In particular, every compact operators on a Hilbert space is the limit of finite rank operators.
Proof. Since $T$ is compact, $S = T^*T$ is a positive compact operator. By the spectral theorem for compact normal operators, let $\{v_i\}$ be the union of the orthonormal bases of $\ker(S - aI)$ for all $a \in \sigma(S) \setminus \{0\}$. So every $v_i$ is the eigenvector of some $a_i \in \sigma(S) \setminus \{0\} \subseteq (0, +\infty)$. If $\sigma(S)$ is infinite, we may arrange the $a_i$’s so they converge to 0.

Let $c_i = \sqrt{a_i}$ and let $e_i = (Tv_i)/c_i$. For $i \neq j$, $(Tv_i, Tv_j) = (Sv_i, v_j) = a_i(v_i, v_j) = 0$. Also, $(Tv_i, Tv_i) = a_i(v_i, v_i) = c_i^2$ implies $\|e_i\| = 1$. Hence, $\{e_i\}$ is an orthonormal set.

For all $x \in H, Tx = \sum_i c_i(x, v_i)e_i$ as the two sides agree on $\text{span}\{v_i\}$ and $\text{span}\{v_i\}^\perp$ (since $Tv_j = c_jv_j$ and $x \perp \text{span}\{v_i\}$ implies $Sx = 0$ and $0 = (Sx, x) = \|Tx\|^2$), which span $H$.

Remarks. (1) The compact self-adjoint case of the spectral theorem is known as the Hilbert-Schmidt theorem.

(2) Let $\{\lambda_j\}$ be the sequence $\{c_i\}$ of nonzero eigenvalues of a compact normal operator $T$ with each $c_i$ repeated $\dim H_{c_i}$ times and $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots$. For each $c_i$, choose an orthonormal basis for $H_{c_i}$. Let the union of these bases form a sequence $\{x_j\}$. Then for every $x \in H, Tx = \sum_i c_i P_{c_i}x = \sum_j \lambda_j(x, x_j)x_j$.

$(\sum \lambda_j(x_j \otimes x_j)$ is called the Schur representation of $T).$ Let $B$ be the union of $\{x_j\}$ and a basis of $H_0 = \ker T$.

Then the matrix of $T$ with respect to $B$ is diagonal. For the sequence $\{\lambda_j\}$, $|\lambda_1| = r(T) = \|T\|$ and

$$|\lambda_{n+1}| = \max\{|(Tx, x) | : x \perp x_1, \ldots, x_n, \|x\| = 1\}$$

because for such $x$, we have $|(Tx, x)| = |\sum_{j \geq n+1} \lambda_j(x, x_j)^2 | \leq |\lambda_{n+1}| \sum_{j \geq n+1} |(x, x_j)|^2 \leq |\lambda_{n+1}||x||^2 = |\lambda_{n+1}|$ and equality is obtained in case $x = x_{n+1}$.

(3) (Schmidt’s Formula) If $\dim H = \infty, T \in L(H)$ is compact normal and $c \not\in \sigma(T)$, then for every $y \in H$, the equation $(cI - T)x = y$ has the solution

$$x = (cI - T)^{-1}y = \frac{1}{c}y + \frac{1}{c} \sum_j \frac{\lambda_j}{c - \lambda_j} (y, x_j)x_j,$$

where $\lambda_j$ and $x_j$ are as in (2) above. To see this, note $\dim H = \infty$ implies $c \neq 0 \in \sigma(T)$. From $cx - Tx = y$, we get

$$x = \frac{1}{c}y + \frac{1}{c}Tx = \frac{1}{c}y + \frac{1}{c} \sum_j \lambda_j(x, x_j)x_j.$$

Taking inner product with $x_k$, we get $(x, x_k) = \frac{1}{c}(y, x_k) + \frac{1}{c} \lambda_k(x, x_k)$. So $(x, x_k) = \frac{1}{c - \lambda_k} (y, x_k)$ and the formula follows.

(4) (Courant’s Minimax Principle) For a compact self-adjoint operator $T \in L(H)$, let $\lambda_1^+ \geq \lambda_2^+ \geq \lambda_3^+ \geq \cdots \geq 0 > \cdots \geq \lambda_2^- \geq \lambda_1^- \geq \cdots$ be the eigenvalues of $T$ in descending order. Then

$$\lambda_n^+ = \inf \{\sup \{(Tx, x) : x \in E_{n-1}^+, \|x\| = 1\} : E_{n-1}^+ \text{ is a } n-1 \text{ dimensional subspace of } H\}$$

$$\lambda_n^- = \sup \{\inf \{(Tx, x) : x \in E_{n-1}^+, \|x\| = 1\} : E_{n-1}^+ \text{ is a } n-1 \text{ dimensional subspace of } H\}.$$