Functional Analysis Oral Exam study notes

Notes transcribed by Mihai Nica

Abstract. These are some study notes that I made while studying for my oral exams on the topic of Functional Analysis. I took these notes from parts of the textbooks "A Course in Functional Analysis" by John B. Conway [**1**], "Functional Analysis" by Peter Lax [**2**] and "Methods of modern mathematical physics: Functional analysis" by Michael Reed, Barry Simon[**3**] and also a very nice real life course taught by Sinan Gunturk in by Spring 2013 at Courant . Please be extremely caution with these notes: they are rough notes and were originally only for me to help me study. They are not complete and likely have errors. I have made them available to help other students on their oral exams.

Contents

Baire Category Theorem

These are based on in class notes and also from [**1**].

Definition. A set A in a metric space is called *nowhere dense* if the closure of A has empty interior. $(\bar{A})^\circ = \emptyset$.

Example. The rationals are NOT nowhere dense. A finite number of points is nowhere dense. Cantor sets are nowhere dense sets. Subsets of nowhere dense sets are nowhere dense.

Remark. Finite unions of nowhere dense sets are still nowhere dense. $\overline{\cup_{i=1}^n A_i}^\circ = \left(\cup_{i=1}^n \bar{A}_i\right)^\circ = \cup_{i=1}^n \bar{A}_i^\circ = \emptyset$.

Definition. A set A which can be written as a countable union of nowhere dense sets is called *1st Category* or *Meagre.* Sets which cannot be written this way are called *2nd Category* or *Nonmeagre.*

Example. The countable union of meagre sets is still meager. Any subset of a meager set is still meager.

Lemma. *Let X be a complete metrix space. If $\{U_n\}$ is a sequence of open dense sets, then $\cap_n U_n$ is also dense.*

Proof. Skip for now. □

Theorem. *A complete metric space is always second catergory or non-meager.*

Proof. Skip for now. □

Important Consequences:

Name	Statement
Banach-Schauder Open Mapping Theorem	Let X, Y be Banach spaces and let $T \in B(X, Y)$ be a bounded linear map. Suppose morevoer that T is onto. Then T is an open map.
Corr	If T is a continuous linear bijection from X to Y then T^{-1} is continous too.
Corr	If $\lVert\cdot\rVert_1$ and $\lVert\cdot\rVert_2$ are two norms on a space X, and there is an m so that $\lVert\cdot\rVert_1 \leq m \lVert\cdot\rVert_2$, then there exists M so that $\lVert\cdot\rVert_1 \geq M \lVert\cdot\rVert_2$
The Closed Graph Theorem	Let X, Y be Banach spaces and let $T : X \to Y$ be linear. Let $\Gamma(T) = \{(x, T(x)) : x \in X\}$ be the graph of T. Then T is continuous if and only if $\Gamma(T)$ is closed.
Banach-Steinhaus Uniform Boundedness Principle	Suppose X, Y are Banach spaces and $(T_\alpha)_{\alpha\in\Lambda}$ is a collection of bounded linear maps. Let $E = \{x \in X : \sup_{\alpha\in\Lambda} \lVert T_\alpha x\rVert < \infty\}$. If E is 2nd category or nonmenager, then $\sup_{\alpha\in\Lambda} \lVert T_\alpha\rVert < \infty$. I.e. the $T_\alpha' s$ are uniformly bounded. By the Baire category theorem, it is enough to show $E = X$.
(Slightly Stronger version)	(Same setup as above) Let $M = E^c = \{x \in X : \sup_\alpha \lVert T_\alpha x\rVert = \infty\}$. Then either M is empty or M is a dense G_δ set.

Theorem. *(Banach-Steinhaus Uniform Boundedness Principle) Suppose X, Y are Banach spaces and $(T_\alpha)_{\alpha\in\Lambda}$ is a collection of bounded linear maps. Let $E = \{x \in X : \sup_{\alpha\in\Lambda} \lVert T_\alpha x\rVert < \infty\}$. If $E = X$ is all of X then $\sup_{\alpha\in\Lambda} \lVert T_\alpha\rVert < \infty$. I.e. the $T_\alpha' s$ are uniformly bounded.*

Proof. Let $E_n = \{x \in X : \sup_{\alpha\in\Lambda} \lVert T_\alpha x\rVert \leq n\}$ so that $E = \cup_n E_n$. Notice also that $E_n = \cap_\alpha \lVert T_\alpha(\cdot)\rVert^{-1} [0, n]$ is an intersection of closed sets (because $x \to \lVert T_\alpha x\rVert$ is continuous), so E_n is closed. Since E is not 1st category, we know that E cannot be written as a countable union of nowhere dense sets. Hence it must be the case that at least one E_n is not nowhere dense. In other words, $\exists n_0$ so that $E_{n_0}^\circ \neq \emptyset$. Hence $\exists x_0, r$ so that $\overline{B_r(x_0)} \subset E_{n_0}$.

For any x with $\lVert x\rVert \leq r$ now, notice that $x_0 + x \in \overline{B_r(x_0)} \subset E_{n_0}$. Hence for such x, we know by definition of E_{n_0} that $\sup_\alpha \lVert T_\alpha(x_0 + x)\rVert \leq n_0$. Have then for any $\lVert x\rVert \leq r$:

$$\begin{aligned}
\sup_\alpha \lVert T_\alpha x\rVert &= \sup_\alpha \lVert T_\alpha(x_0 + x) - T_\alpha(x_0)\rVert \\
&\leq \sup_\alpha \left(\lVert T_\alpha(x_0 + x)\rVert + \lVert T_\alpha(x_0)\rVert\right) \\
&\leq n_0 + n_0 = 2n_0
\end{aligned}$$

So by scaling, we conclude that for any x with $\lVert x\rVert \leq 1$ that $\sup_\alpha \lVert T_\alpha x\rVert \leq \frac{2n_0}{r}$. Have finally then that $\sup_\alpha \lVert T_\alpha\rVert = \sup_\alpha \sup_{\lVert x\rVert=1} \lVert T_\alpha x\rVert \leq \frac{2n_0}{r} < \infty$. □

Theorem. *(The slightly stronger version) (Same setup as Banach-Steinhaus) Let $M = E^c = \{x \in X : \sup_\alpha \lVert T_\alpha x\rVert < \infty\}$. Then either M is empty or M is a dense G_δ set.*

Let $U_n = \{x \in X : \sup_\alpha \|T_\alpha x\| > n\}$ so that $M = \bigcap_\alpha U_n$. Notice that we can write:

$$\begin{aligned} U_n &= \bigcup_\alpha \{x \in X : \|T_\alpha x\| > n\} \\ &= \bigcup_\alpha \left(\|T_\alpha(\cdot)\|^{-1}(n, \infty) \right) \end{aligned}$$

since the map $x \to \|T_\alpha x\|$ is continuous each set $\|T_\alpha(\cdot)\|^{-1}(n,\infty)$ is open and we see from this that U_n is a union of open sets. Since $M = \bigcap_\alpha U_n$, we see that M is a G_δ-set.

Claim: Either M is empty or U_n is dense set for every $n \in \mathbb{N}$.

Pf: It suffices to show the following: if there is a single n_0 for which $U_{n_0}$is not dense, then M is empty. Suppose $U_{n_0}$is not dense. Then, by definition of dense, $\overline{U_{n_0}} \neq X$. In other words this is $\overline{U_{n_0}}^c \neq \emptyset$. Now, $\overline{U_{n_0}}$ is a closed set, so we know that $\overline{U_{n_0}}^c$ is an open set. Hence, since this is a non-empty open set, we can find $x_0 \in X$ and $r > 0$ so that $\overline{B_r(x_0)} \subset \overline{U_{n_0}}^c$.

Consider any $x \in X$ with $\|x\| \leq r$. Then $x_0 + x \in \overline{B_r(x_0)} \subset \overline{U_{n_0}}^c$. Hence $x_0 + x \notin U_{n_0}$. By definition of U_n this means that $\sup_\alpha \|T_\alpha(x_0 + x)\| \leq n$. Using scaling and translation invariance, we have then that for any x with $\|x\| \leq 1$ that:

$$\begin{aligned} \sup_\alpha \|T_\alpha x\| &= \frac{1}{r} \sup_\alpha \|T_\alpha(rx)\| \\ &= \frac{1}{r} \sup_\alpha \|T_\alpha(x_0 + rx) - T_\alpha(x_0)\| \\ &\leq \frac{1}{r} \sup_\alpha \left(\|T_\alpha(x_0 + rx)\| + \|T_\alpha(x_0 + 0)\|\right) \\ &\leq \frac{1}{r}(n_0 + n_0) \text{ since } \|rx\| \leq r \text{and } \|0\| \leq r \\ &= \frac{2n_0}{r} \end{aligned}$$

Finally then we see that the $T_\alpha$are uniformly bounded,

$$\begin{aligned} \sup_\alpha \|T_\alpha\| &= \sup_\alpha \sup_{\|x\|=1} \|T_\alpha x\| \\ &\leq \frac{2n_0}{r} < \infty \end{aligned}$$

This means that M is the empty set, because for every $x \in X$ we have that $\|T_\alpha x\| \leq \sup_\alpha \|T_\alpha\| \|x\| < \infty$ so $x \notin M$. □

Combining the initial remarks and the claim we see that M is either empty or otherwise we have that $M = \cap_n U_n$ and every U_n is dense. Since the countable intersection of open dense sets is dense (this was the main lemma in the pf of Baire's theorem), in the latter case we see that M is a dense G_δ set, as desired. ■

The Hahn-Banach Theorem

These are based on in class notes and also from [**1**].
"H-B" = "Hahn-Banach" for the rest of this section.
In all the statements the set up is:

$$\begin{aligned}
(X, \|\cdot\|) &\equiv \text{A normed vector space over the field } \mathbb{F} \\
\mathbb{F} &\equiv \text{The field that } X \text{is over. Will either be } \mathbb{R} \text{or } \mathbb{C} \\
M &\equiv \text{A linear subspace of } X \\
p &\equiv \text{A sub-linear functional } p : X \to \mathbb{R}, \text{ i.e. } p \text{satisfies:} \\
&\quad\; p(x+y) \le p(x) + p(y),\ p(ax) = ap(x)\ \forall a > 0. \\
q &\equiv \text{A semi-norm } q : X \to \mathbb{R}, \text{ i.e. } q \text{satisfies:} \\
&\quad\; q(x+y) \le q(x) + q(y),\ q(\lambda x) = |\lambda| q(x)\ \forall \lambda \in \mathbb{C}. \\
&\quad\; \text{(Rmk: semi-norm is stricter than sub-linear functional)} \\
\ell_A &\equiv \text{A linear functional } \ell_A : A \to \mathbb{F} \text{where } A \text{will be some subspace of } X \\
\ell_A \le p &\equiv \text{Shorthhand for:} \ell_A(x) \le p(x)\ \forall x \in A \\
\ell_A \le q &\equiv \text{Shorthhand for:} \ell_A(x) \le q(x)\ \forall x \in A \\
\ell_B \overset{ext.}{\succeq} \ell_A &\equiv \text{"}\ell_B \text{ extends } \ell_A\text{", shorthand for " } A \subset B \text{and } \ell_A(x) = \ell_B(x) \forall x \in A\text{"}
\end{aligned}$$

Below is a table with all the different "flavours" of the H-B theorem.

Name	$\mathbb{F}$	Hypothesis	Conclusion	"Pf"
Baby H-B Thm	$\mathbb{R}$	$\ell_M \leq p$; $x_0 \in X - M$. Define $M \oplus x_0\mathbb{R} = \{x + \lambda x_0 : x \in M, \lambda \in \mathbb{R}\}$	$\exists \ell_{M\oplus x_0\mathbb{R}} \overset{ext.}{\succeq} \ell_M$ so that $\ell_{M\oplus x_0\mathbb{R}} \leq p$	Can find the correct $\ell_{M\oplus x_0\mathbb{R}}$as long as we can define $\ell_{M\oplus x_0\mathbb{R}}(x_0)$ to be in some particular interval. The fact $\ell_M \leq p$ can be used to show that this interval is non-empty.
Real H-B Thm	$\mathbb{R}$	$\ell_M \leq p$	$\exists \ell_X \overset{ext.}{\succeq} \ell_M$ so that $\ell_X \leq p$	Zorn's Lemma is used with the partial ordering "$\overset{ext.}{\succeq}$". Every totally ordered set has a maximal element by taking the union of all the subspaces. By Zorn's Lemma, there is a maximal element. By the Baby H-B Thm, the maximal element cannot be a proper subspace of X.
Complex H-B Thm	$\mathbb{C}$	$\lvert\ell_M\rvert \leq q$	$\exists \ell_X \overset{ext.}{\succeq} \ell_M$so that $\lvert\ell_X\rvert \leq q$	Use the Real H-B Thm to prove this one. Its a pretty unenlightening proof involving manipulations with complex numbers.
(Not named)	$\mathbb{F}$	$x_0 \in X - \{\vec{0}\}$	$\exists \ell \in X^\star$s.t. $\lVert\ell\rVert_{X^\star} = 1$,$\ell(x_0) = \lVert x\rVert$	Apply the H-B thm on the space $M = \{\lambda x_0 : \lambda \in \mathbb{F}\}$ with the functional $\ell_M(\lambda x_0) := \lambda \lVert x_0\rVert$ and seminorm $q(x) = \lVert x\rVert$. The extension $\ell_X \overset{ext.}{\succeq} \ell_M$ is what we want. The ineq $\lvert\ell_X\rvert \leq q$ gives that $\lVert\ell\rVert_{X^\star} \leq 1$ by linearity. The other inequality is clear by pluggin in x_0.
Analytic H-B Thm	$\mathbb{F}$	$\ell_M \in M^\star$	$\exists \ell_X \in X^\star$, $\ell_X \overset{ext.}{\succeq} \ell_M$ with $\lVert\ell_X\rVert_{X^\star} = \lVert\ell_M\rVert_{M^\star}$	Let $q(x) = \lVert\ell_M\rVert_{M^\star} \lVert x\rVert$ be the seminorm. Then apply H-B thm. The ineq $\lvert\ell_X\rvert \leq q$ gives that $\lVert\ell_X\rVert_{X^\star} \leq \lVert\ell_M\rVert_{M^\star}$ by linearity. The other ineq is clear since $M \subset X$
Projection H-B Thm	$\mathbb{F}$	$x_0 \in X - \bar{M}$ such that $\text{dist}(x_0, M) > 0$	$\exists \ell \in X^\star$s.t. $\ell\rvert_M = 0$ and $\ell(x_0) = 1$ and $\lVert\ell\rVert_{X^\star} = \frac{1}{\text{dist}(x_0,M)}$	Let $M_1 = M \oplus x_0\mathbb{F}$ and define $\ell_{M_1}(x + \lambda x_0) = \lambda$. $\lVert\ell_{M_1}\rVert = \sup_{x\in M,\lambda\in\mathbb{F}} \frac{\lvert\lambda\rvert}{\lVert x+\lambda x_0\rVert} = \sup_{x\in M,\lambda\in\mathbb{F}} \frac{1}{\lVert\frac{x}{\lambda}+x_0\rVert} = \frac{1}{\text{dist}(x_0,M)}$ Then use the Analytic H-B thm to extend to ℓ_X.

THEOREM. *(Baby H-B Thm) Suppose*$\ell_M \leq p$ *and that* $x_0 \in X - M$*. Define* $M \oplus x_0\mathbb{R} = \{x + \lambda x_0 : x \in M, \lambda \in \mathbb{R}\}$*. Then* $\exists \ell_{M\oplus x_0\mathbb{R}} \overset{ext.}{\succeq} \ell_M$ *so that* $\ell_{M\oplus x_0\mathbb{R}} \leq p$*.*

PROOF. Suppose we found a value$\ell(x_0)$ that we liked a lot. Then we could define: $\ell_{M\oplus x_0\mathbb{R}}(x+\lambda x_0) = \ell_M(x)+\lambda\ell(x_0)$ and we would have found the functional $\ell_{M\oplus x_0\mathbb{R}} \overset{ext.}{\succeq} \ell_M$ we want! Of course, since we want $\ell_{M\oplus x\mathbb{R}} \leq p$, not just any value of $\ell(x_0)$ will do. We need $\ell(x_0)$ to obey the following inequalities:

Claim 1: For a fixed value $\ell(x_0)$, define $\ell_{M\oplus x_0\mathbb{R}}(x + \lambda x_0) = \ell_M(x) + \lambda\ell(x_0)$. Then:

$$\ell_{M\oplus x_0\mathbb{R}} \leq p \iff \forall x \in M, \begin{array}{l} \ell_M(x) + \ell(x_0) \leq p(x + x_0) \quad \text{and} \\ \ell_M(x) - \ell(x_0) \leq p(x - x_0) \end{array}$$

Pf: ($\Rightarrow$) Plug in $x \pm x_0$ into $\ell_{M\oplus x_0\mathbb{R}} \leq p$, get $\ell_{M\oplus x_0\mathbb{R}}(x \pm x_0) \leq p(x \pm x_0)$. By using the definition of $\ell_{M\oplus x_0\mathbb{R}}$ on the LHS, we get the desired inequalities.

($\Leftarrow$)Let $x + \lambda x_0 \in M \oplus x_0\mathbb{R}$ be arbitrary. There are two cases, one where $\lambda > 0$ and one where $\lambda \leq 0$. We handle both cases simultaneously by using the $\pm$sign abusivly and writing $\lambda = \pm|\lambda|$. Write:

$$\begin{aligned}
\ell_{M\oplus x_0\mathbb{R}}(x+\lambda x_0) &= \ell_{M\oplus x_0\mathbb{R}}(x \pm |\lambda|\, x_0) \\
&= |\lambda| \left(\ell_{M\oplus x_0\mathbb{R}} \left(\frac{x}{|\lambda|} \pm x_0 \right) \right) \\
&= |\lambda| \left(\ell_M \left(\frac{x}{|\lambda|} \right) \pm \ell\,(x_0) \right) \text{ by def'n of } \ell_{M\oplus x_0\mathbb{R}} \\
&\le |\lambda| \left(p \left(\frac{x}{|\lambda|} \right) \pm p(x_0) \right) \text{ by the hypothesis inequalities} \\
&= p\,(x \pm |\lambda|\, x_0) \text{ since } p \text{is a sublinear functional} \\
&= p(x+\lambda x_0)
\end{aligned}$$

So indeed, $\ell_{M\oplus x_0\mathbb{R}} \le p$ □

To show that a value of $\ell(x_0)$ exists which satisfies the inequalities from Claim 1, we need the following to hold for all $x \in M$:

$$\ell_M(x) - p(x - x_0) \le \ell(x_0) \le p(x + x_0) - \ell_M(x)$$

It suffices then to show that $\forall x_1, x_2 \in M$ that $\ell_M(x_1) - p(x_1 - x_0) \le p(x_2 + x_0) - \ell_M(x_2)$. Indeed, this is a consequence of $\ell_M \le p$. Pluggin in $x_1 + x_2$ into $\ell_M \le p$, we have:

$$\begin{aligned}
\ell_M(x_1) + \ell_M(x_2) &= \ell_M(x_1 + x_2) \\
&\le p(x_1 + x_2) \\
&= p\,((x_1 - x_0) + (x_0 + x_2)) \\
&\le p(x_1 - x_0) + p(x_0 + x_2)
\end{aligned}$$

Rearranging now gives the desired inequality. □

LEMMA. *(Zorn's Lemma) A* partial ordering *on a set P is a relation " $\preceq$ " that is reflexive ($a \preceq a$), antisymetric ($a \preceq b, b \preceq a \implies a = b$), and transitive ($a \preceq b, b \preceq c \implies a \preceq c$). Suppose that every totally ordered subset (i.e. a set in which for every pair a, b either $a \preceq b$ or $b \preceq a$), $\{a_\alpha\}_{\alpha\in\Lambda}$ has an upper bound in P (i.e. an element $a_{\star,\Lambda} \in P$ so that $a_\alpha \preceq a_{\star,\Lambda}$for all $\alpha \in \Lambda$). Then P contains at least one maximal element (i.e. an $a_\star$ so that $a \preceq a_\star$ for all $a \in P$).*

REMARK. This is equivalent to the axiom of choice, but the proof is non-trivial!

THEOREM. *(Real H-B Theorem) Let X be a normed vector space over $\mathbb{R}$, p a sublinear functional on X, M a subspace, and $\ell_M : M \to \mathbb{R}$ a linear functional such that $\ell_M \le p$ (i.e.$\ell_M(x) \le p(x)\,\forall x \in M$). Then $\exists \ell_X$ a linear functional that extends $\ell_X \overset{ext.}{\succeq} \ell_M$ (i.e. $\ell_M(x) = \ell_X(x) \forall x \in M$) and $\ell_X \le p$ (i.e.$\ell_X(x) \le p(x)\,\forall x \in X$)*

PROOF. Let $P = \left\{ \ell_A : A \to \mathbb{R} : \ell_A \overset{ext.}{\succeq} \ell_M \right\}$ be the space of all linear functions which are defined on subspaces A of X. Then " $\overset{ext.}{\succeq}$ " is a partial ordering on P (Rmk: one way to see this is to notice that $\overset{ext.}{\succeq}$ is inclusion of the graphs, that is $\ell_A \overset{ext.}{\succeq} \ell_B$iff Graph $(\ell_A) \supset$ Graph (ℓ_B) where Graph$(f) = \{(x, f(x)), x \in \text{Domain }(f)\}$. Moreover, every totally ordered subset has a maximum element in P. Namely, if $\{\ell_{A_\alpha}\}_{\alpha\in\Lambda}$ is a totally ordered set, then define $A_{\star,\Lambda} = \cup_{\alpha\in\Lambda} A_\alpha$ and $\ell_{A_{\star,\Lambda}}(x) = \ell_{A_\alpha}(x)$ for $x \in A_\alpha$. (This is well defined because $\{\ell_{A_\alpha}\}_{\alpha\in\Lambda}$ is a totally ordered set).

Now by the conclusion of Zorn's lemma, there is a maximal element $\ell_{A_\star}$ for all of P. Now we claim that $A_\star = X$. Indeed, if by contradiction, $A_\star \neq X$, then there is at least one element $x_0 \in X - A_\star$. But now by the Baby H-B Thm, we can get an extension $\ell_{A_\star \oplus x_0\mathbb{R}}$. But this contradicts the maximality of $\ell_{A_\star}$in P! So it must be that $A_\star = X$. □

Theorem. *(Complex H-B Theorem) Let X be a normed vector space over $\mathbb{C}$, q a seminorm on X, M a subspace, and $\ell_M : M \to \mathbb{C}$ a linear functional such that $|\ell_M| \leq q$ (i.e.$|\ell_M(x)| \leq q(x)\ \forall x \in M$). Then $\exists \ell_X$ a linear functional that extends $\ell_X \overset{ext.}{\succeq} \ell_M$ (i.e. $\ell_M(x) = \ell_X(x) \forall x \in M$) and $|\ell_X| \leq q$ (i.e.$|\ell_X(x)| \leq q(x)\ \forall x \in X$)*

Proof. (By manipulations using the Real H-B Thm) Let $u_M(x) = \mathrm{Re}(\ell_M(x))$ and $v(x) = \mathrm{Im}(\ell_M(x))$ so that $\ell_M = u_M + iv_M$. u_M and v_M are seen to be $\mathbb{R}-$linear functionals, because ℓ_M is $\mathbb{R}$-linear. (ℓ_M is more than $\mathbb{R}-$linear actually!) Since ℓ_M is actually $\mathbb{C}$-linear, we have that:

$$v_M(x) = \mathrm{Im}\ (\ell_M(x)) = \mathrm{Re}(-i\ell_M(x)) = \mathrm{Re}(\ell_M(ix)) = u_M(ix)$$

So then $\ell_M(x) = u_M(x) + iu_M(ix)$ can be entirely reconstructed from u_M. Now, q being a semi-norm, is also a sublinear map (which is a slightly looser condition), and $u_M(x) \leq |\ell_M(x)| \leq q(x)$. So applying the Real H-B Thm we get a $u_X \overset{ext.}{\succeq} u_M$ and $u_X \leq q(x)$. Now let $\ell_X(x) = u_X(x) + iu_X(ix)$. One now verifies that $\ell_X \overset{ext.}{\succeq} \ell_M$ (our calculation early basically did this). Finally to check that $|\ell_X| \leq q$ have:

$$\begin{aligned}
|\ell_X(x)| &= e^{i\theta}\ell_X(x) \text{ for some } \theta \\
&= \ell_X(e^{i\theta}x) \\
&= \mathrm{Re}(\ell_X(e^{i\theta}x)) \text{ since the LHS is real} \\
&= u_X(e^{i\theta}x) \\
&\leq q(e^{i\theta}x) \text{ since } |\ell_X| \leq q \\
&= |e^{i\theta}|q(x) = q(x) \text{ since } q\text{is a seminorm.}
\end{aligned}$$

□

Hilbert Spaces

These are notes from Chapter 1 of [**1**].

3.1. Elementary Properties and Examples

Definition. 1.1. Definition of a **semi-inner produc**t on a vector space $\mathscr{X}$; it is sesqui-linear (i.e its bilinear except for conjugation in the second slot), non-negative definite (i.e. $\langle x, x\rangle \geq 0$) and "Hermitian" ($\langle x, y\rangle = \overline{\langle y, x\rangle}$). An **inner product** is one that is also positive definite, $\langle x, x\rangle = 0 \iff x = 0$.

Example. 1.2-1.3 $L^2(\mu)$ on a measure space (X, Ω, μ).

Theorem. *1.4 (Cauchy-Schwarz Ineq) If* $\langle\cdot,\cdot\rangle$*is a semi-inner product on* $\mathscr{X}$ *then:*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \langle y, y\rangle = \|x\|^2 \|y\|^2$$

Proof. Use $0 \leq \langle x - \alpha y, x - \alpha y\rangle$ and put $\alpha = te^{-i\theta}$ where θ is such that $\langle x, y\rangle = be^{i\theta}$, and get a quadratic in t which has no real roots so its discriminet is non-positive. □

Corollary. *(1.5) a)* $\|x + y\| \leq \|x\| + \|y\|$*, b)* $\|\alpha x\| = |\alpha| \|x\|$ *c)* $\|x\| = 0 \iff x = 0$ *for true inner products.*

Proof. a) follows since:

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\,|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \text{ by C.S.} \end{aligned}$$

. b), c) are straight from the definitions of $\|\cdot\|^2 = \langle\cdot,\cdot\rangle$ and the propoerties of (semi-)inner products □

Definition. (1.6) A Hilbert space is an inner product space where the norm $\|\cdot\|$ induced by the inner product yields a complete space (i.e. Cauchy sequences converge)

Example. (1.7) $L^2(\mu)$ or $\ell^2(I)$ for any set I. To be complete in our presentation, we would have to prove these are complete. To do this for $\ell^2(I)$, you observe that for any Cauchy sequence, the indivdual coordinates are Cauchy. Hence we are converging coordinatwise to something. You then do some estimates to verifty that this coordinate-wise limit is in ℓ^2 and that the sequence actually converges to this,

In $L^2(\mu)$ you can use the fact that if a sequence of function is Cauchy, then you can find a subsequence where $\mathbf{P}\left(\left|f_{n_k} - f_{n_{k+1}}\right| > \frac{1}{2^k}\right) < \frac{1}{2^k}$ and so by Borel Cantelli/facts about types of convergence, this subsequence is a.e. Cauchy and hence converges almost everywhere to something. Again, verify this a.e. limit is in

L^2 and that we in fact converge to it in L^2. Finally, for Cauchy sequences, if one subsequence converges, the whole sequence must coverge too.

This would be worth doing in full detail sometime....but I don't feel like it now.

PROPOSITION. *(1.9) (Roughly) The completion of an inner product space is a Hilbert space: that is if $\mathscr{X}$ is an incomplete inner product space with inner product $\langle\cdot,\cdot\rangle$, then if we let $\mathscr{H}$ be the comletion of $\mathscr{X}$, then the inner product $\langle\cdot,\cdot\rangle$ extends to all of $\mathscr{H}$ in such a way to make $\mathscr{H}$ a Hilbert space.*

I am going to skip some stuff here.... mostly the proof that the "Bergman Space" $L^2_a(G)$ the space of analytic functions on a subset $G \subset \mathbb{C}$ which are square integrable (with respect to area).

3.2. Orthogonality

DEFINITION. (2.1) We say $f \perp g$ (read as: "orthogonal") if $\langle f, g\rangle = 0$ and for subsets $A \perp B$ if $f \perp g \forall f \in A, g \in B$.

THEOREM. *(2.2.) (Pythagoras). If $f_1, f_2, \ldots f_n$ are pairwise orthogonal, then* $\|f_1 + \ldots + f_n\|^2 = \|f_1\|^2 + \ldots + \|f_n\|^2$

PROOF. Exapand out $\langle f_1 + \ldots + f_n, f_1 + \ldots + f_n\rangle$ and use $f_i \perp f_j$. (Or to see more precisly: use induction) □

THEOREM. *(2.3) (Parallelogram Law) If $\mathscr{H}$ is a Hilbert space, and $f, g \in \mathscr{H}$ then:*

$$\|f+g\|^2 + \|f-g\|^2 = 2\left(\|f\|^2 + \|g\|^2\right)$$

PROOF. Again, just write it as inner product and expand. □

REMARK. The converse is also true: if $\mathscr{H}$ is a normed space that has the parralelogram law, then in fact it is a Hilbert space with inner product defined by: $\langle f, g\rangle = \frac{1}{4}\|f+g\|^2 - \frac{1}{4}\|f-g\|^2$.

REMARK. We will mostly be using this with:

$$\left\|\frac{f+g}{2}\right\|^2 + \left\|\frac{f-g}{2}\right\|^2 = \frac{1}{2}\left(\|f\|^2 + \|g\|^2\right)$$

3.2.1. Projections and Stuff!

DEFINITION. (2.4.) A set A is called convex if $tx + (1-t)y \in A$ for all $x, y \in A$ $t \in (0,1)$

THEOREM. *(2.5.) If $\mathscr{H}$ is a Hilbert space, K a closed convex non-empty subset of $\mathscr{H}$ and $h \in \mathscr{H}$, then there exists a **unique** point $k_0 \in K$ such that:*

$$\|h - k_0\| = \mathit{dist}(h, K) := \inf\{\|h-k\| : k \in K\}$$

REMARK. If we knew that K was <u>compact</u> then the existence statement here would just be the extreme value theorem. Unfortunately, we cant assume this; the unit ball is not even compact here!

PROOF. WOLOG by translating, assume that $h = 0$. Take any sequence $k_n \in K$ so that $\|k_n\| \to d := \mathrm{dist}(0, K) = \inf\{\|k\| : k \in K\}$. The Parrallelogram law and the fact that K is convex will show us that actually k_n is Cauchy.

Have by parallelogram law:

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

Since K is convex, we have that $\frac{1}{2}(k_n+k_m)\in K$ and consequently, $\left\|\frac{1}{2}(k_n+k_m)\right\|^2\geq d^2$ since d is the inf of all points from K. Now, since $\|k_n\|\to d$, for any $\epsilon>0$ we may choose N so large so that $n>N\implies\|k_n\|^2<d^2+\frac{1}{4}\epsilon^2$ and we see then that for any $n,m>N$ that:

$$\left\|\frac{k_n-k_m}{2}\right\|^2<\frac{1}{2}\left(d^2+\frac{1}{4}\epsilon^2+d^2+\frac{1}{4}\epsilon^2\right)-d^2=\frac{1}{4}\epsilon^2$$

And so we see that k_n is a Cauchy sequence! Since K is closed and $\mathscr{H}$ is complete, we have a limit point k_0 of the sequence k_n. Continuity of $\|\cdot\|$ shows that $\|k_0\|=\lim_{n\to\infty}\|k_n\|=d$ by our choice!

To prove uniqueness we again use that K is convex . If $k_0$and $h_0\in K$ are two points that minizize $\|\cdot\|$, then by convexity $\frac{1}{2}\left(k_0+h_0\right)\in K$ too, and hence:

$$d\leq\left\|\frac{1}{2}(h_0+k_0)\right\|\leq\frac{1}{2}\left(\|h_0\|+\|k_0\|\right)=d$$

So $\frac{1}{2}(h_0+k_0)$ is a minimizer too! But then by Parallleogram law we have:

$$d^2=\left\|\frac{h_0+k_0}{2}\right\|^2=d^2-\left\|\frac{h_0-k_0}{2}\right\|^2$$

Shows $h_0=k_0$. □

Theorem. *(2.6.) If in addition to being closed and convex a set $\mathscr{M}$ is a closed linear subset of $\mathscr{H}$. Let $h\in\mathscr{H}$. Have:*

$$\begin{aligned}\|h-f_0\| &= dist(h,\mathscr{M}) \qquad\iff h-f_0\perp\mathscr{M}\\ &= \inf\{\|h-f\|:f\in\mathscr{M}\}\end{aligned}$$

Proof. ($\Rightarrow$) Suppose $f_0\in\mathscr{M}$ and $\|h-f_0\|=\text{dist}\left(h,\mathscr{M}\right)$. Then $f_0+f\in\mathscr{M}$ for all $f\in\mathscr{M}$ and we have:

$$\begin{aligned}\|h-f_0\|^2 &\leq \|h-(f+f_0)\|^2\\ &= \|h-f_0\|^2-2\text{Re}\langle h-f_0,f\rangle+\|f\|^2\end{aligned}$$

Thus:

$$2\text{Re}\langle h-f_0,f\rangle\leq\|f\|^2$$

This holds for any $f\in\mathscr{M}$. Now the LHS $\to 0$ linearly in $\|f\|$while the RHS $\to 0$ quadratically, so this can only work if the LHS is 0. To make this more precise: Let r,θ so that $\langle h-f_0,f\rangle=re^{i\theta}$ and plug in $g=te^{i\theta}$ into $2\text{Re}\langle h-f_0,g\rangle\leq\|g\|^2$ to get $2tr\leq t^2\left\|f^2\right\|$. This inequality can only hold for every t if $r=0$!

($\Leftarrow$) Suppose $h-f_0\perp\mathscr{M}$. Then for any $f\in\mathscr{M}$ we have $h-f_0\perp f-f_0$ this gives:

$$\begin{aligned}\|h-f\|^2 &= \|h-f_0\|^2+\|f-f_0\|^2 \text{ by Pythag}\\ &\geq \|h-f_0\|^2\end{aligned}$$

So indeed, f_0 is the minimizer!! □

DEFINITION. If $A \subset \mathscr{H}$ then define:

$$A^{\perp} = \{f \in \mathscr{H} : f \perp g \forall g \in A\}$$

REMARK. For any set A, the orthogonal space $A^{\perp}$is always a closed linear subspace of $\mathscr{H}$.

DEFINITION. The above theorems show that if $\mathscr{M}$ is a closed linear subspace of $\mathscr{H}$ and $h \in \mathscr{H}$,then there is a unique element f_0 in $\mathscr{M}$ such that $h - f_0 \in \mathscr{M}^{\perp}$(its the same f_0 that minimizes $\|h - f_0\|$). Define $P : \mathscr{H} \to \mathscr{M}$ by $Ph = f_0$

THEOREM. *(2.7) If $\mathscr{M}$ is a closed linear subspace of $\mathscr{H}$ and $h \in \mathscr{H}$, define Ph to be the unique point in $\mathscr{M}$ such that $h - Ph \perp \mathscr{M}$ then:*
a) P is a linear transformation.
b) $\|Ph\| \leq \|h\|$
c)$P^2 = P$
d) $\ker P = \mathscr{M}^{\perp}$and $ranP = \mathscr{M}$

PROOF. By our previous theorems, we have two characterizations of Ph, firstly that $h - Ph \perp \mathscr{M}$ and secondly that $\|h - Ph\|$ is minimal. Sometimes it is easier to use one characterization and sometimes it is easier to use another! (Actually I mostly used the first one)

a) $h_1 - Ph_1 \perp \mathscr{M}$ and $h_2 - Ph_2 \perp \mathscr{M}$, so by linearity of $\langle \cdot, \cdot \rangle$ we know that $(h_1 + \alpha h_2) - (Ph_1 + \alpha Ph_2) \perp \mathscr{M}$. But $P(h_1 + \alpha h_2)$ is the unique element of $\mathscr{M}$ which has $(h_1 + \alpha h_2) - P(h_1 + \alpha h_2) \perp \mathscr{M}$. so we conclude that $(Ph_1 + \alpha Ph_2) = P(h_1 + \alpha h_2)$

b) $\|h\|^2 = \|h - Ph\|^2 + \|Ph\|^2$ by Pythagoras and the result follows.

c) $Ph - Ph = 0$ so it is true that $Ph - Ph \perp \mathscr{M}$. Hence $P(Ph) = Ph$ again by the uniqueness.

d) Any $f \in \ker P$ has $f - 0 \perp \mathscr{M}$ so $f \in \mathscr{M}^{\perp}$by definition. The reverse inclusions follows by the uniqueness. $\text{ran}P = \mathscr{M}$ becauseOrt $Pf = f$ for every $f \in \mathscr{M}$ and $Pf \in \mathscr{M}$ is always true by def'n of P. □

DEFINITION. This P is claeed the **orthogonal projection**.

DEFINITION. Write $\mathscr{M} \leq \mathscr{H}$ to mean that $\mathscr{M}$ is a closed linear subspace of $\mathscr{H}$.

DEFINITION. We say $\mathscr{Y}$ is a linear manifold if it is a linear subspace which is not nessisarily closed.

COROLLARY. *(2.9) If $\mathscr{M}$ is a closed linear subspace of $\mathscr{H}$ then $\left(\mathscr{M}^{\perp}\right)^{\perp} = \mathscr{M}$*

PROOF. By an easy exercise, if $\mathscr{M}$ is a linear subspace then we can write $I = P_{\mathscr{M}} + P_{\mathscr{M}^{\perp}}$. Have that $(\mathscr{M}^{\perp})^{\perp} = \ker(P_{M^{\perp}}) = \ker(I - P_{\mathscr{M}}) = \mathscr{M}$ since $(I - P_{\mathscr{M}})f = 0$ if and only if $f \in \mathscr{M}$ (this last fact can be seen by pythagoras for example) □

COROLLARY. *(2.10) If $A \subset \mathscr{H}$ is some subset, then $(A^{\perp})^{\perp} = \overline{span\ \{A\}}$is the closed linear span of A in $\mathscr{H}$.*

PROOF. $\overline{\text{span }\{A\}}$ is a closed linear subspace of $\mathscr{H}$. Hence $\overline{\text{span }\{A\}} = \left(\overline{\text{span }\{A\}^{\perp}}\right)^{\perp}$. But $\overline{\text{span}(A)} = A^{\perp}$by linearity of $\langle \cdot, \cdot \rangle$. Indeed, $\langle f, a \rangle = 0 \forall a \in A \iff \langle f, g \rangle = 0 \forall g \in \overline{\text{span}(A)}$. □

COROLLARY. *(2.11) If $\mathscr{Y}$ is a linear manifold (i.e. a linear subspace which is not nessisarily closed) then $\mathscr{Y}$ is dense in $\mathscr{H}$ if and only if $\mathscr{Y}^{\perp} = \{0\}$.*

PROOF. Have $\mathscr{Y} = \text{span}\,\{\mathscr{Y}\}$ since $\mathscr{Y}$ is a linear manifold, and so by previous corollary, $\left(\mathscr{Y}^{\perp}\right)^{\perp} = \overline{\text{span}(\mathscr{Y})} = \overline{\mathscr{Y}}$.

($\Leftarrow$) If $\mathscr{Y}^{\perp} = \{0\}$ then have $\overline{\mathscr{Y}} = \left(\mathscr{Y}^{\perp}\right)^{\perp} = \{0\}^{\perp} = \mathscr{M}$

($\Rightarrow$) Both sides are closed linear subspaces, so taking $\perp' s$ and using that $\mathscr{M}^{\perp^{\perp}} = \mathscr{M}$ for closed linear subsapces, we have that $\mathscr{Y}^{\perp} = \overline{\mathscr{Y}}^{\perp} = \mathscr{M}^{\perp} = \{0\}$ □

3.3. Riesz Representation Theorem

PROPOSITION. *(3.1) Let $\mathscr{H}$ be a Hilbert space and let $L : \mathscr{H} \to \mathbb{F}$ be a linear functional. The following are equivalent:*

a) L is continuous

b) L is continuous at 0.

c) L is continuous at some point.

d) There is a constant $c > 0$ such that $|L(h)| \le c\,\|h\|$ for every $h \in \mathscr{H}$

PROOF. It is clear that $a \implies b \implies c$ and $d \implies b$. We will show that $c \implies a$ and that $b \implies d$.

($c \implies a$) This is essentially because L is "translation invariant" because it is linear. Say L is continuous at some h_0. To check that L is continuous, take any convergent sequence $h_n \to h$. Then $h_n - h + h_0 \to h_0$ so by continuity of L at h_0 we have that $L(h_n - h + h_0) \to L(h_0)$ using lineariyt of L and rearranging gives the desired result.

($b \implies d$) By continuity, $L^{-1}\,(-1,1)$ is an open set. Since this contains 0, we can find a ball $B_{\delta}(0) \subset L^{-1}(-1,1)$. In other words, $\|h\| < \delta \implies |L(h)| \le 1$. Now for arbitary h, scale h down by a factor of $\delta\left(\|h\| + \epsilon\right)^{-1}$ an apply this to get that $|L(h)| \le \delta^{-1}\left(\|h\| + \epsilon\right)^{-1}$. Since this works for any $\epsilon > 0$, we get the conlusion d) with $c = \delta^{-1}$. □

DEFINITION. (3.2) Such a functional is called a **bounded linear functional.** and we define its norm:

$$\|L\| = \sup\left\{|L(h)| : \|h\| \le 1\right\}$$

THEOREM. *(3.4.) The Riesz Representation Theorem*

If $L : \mathscr{H} \to \mathbb{F}$ is a bounded linear functional, there there is a unique vector $h_0 \in \mathscr{H}$ such that $L(h) = \langle h, h_0\rangle$ for every $h \in \mathscr{H}$. Moreover, $\|L\| = \|h_0\|$

REMARK. The proof uses the theory of orthogonal projections just developed! The vector h_0 must be in $\ker L^{\perp}$ and indeed choosing the right vector from this space gives the result.

PROOF. Let $\mathscr{M} = \ker L$. Because L is continuous, this is a closed linear subspace of $\mathscr{H}$. If L is identiacally 0 then the result is trivial, and otherwise we may find a vector $f_0 \notin \mathscr{M}$. By taking the orthogonal projection onto $\mathscr{M}^{\perp}$, we may assume WOLOG that $f_0 \in \mathscr{M}^{\perp}$. By scaling f_0 we may also assume WOLOG that $L(f_0) = 1$.

The main observation is now that $L\left(h-L(h)f_0\right)=0$ for any $h\in\mathscr{H}$. Hence $h-L(h)f_0\in\mathscr{M}$ and so we have:

$$\begin{aligned} 0 &= \langle h-L(h)f_0,f_0\rangle \\ \implies \langle h,f_0\rangle &= L(h)\left\|f_0\right\|^2 \end{aligned}$$

Let $h_0=\left\|f_0\right\|^{-2}f_0$ now seals the deal.

Uniquness follows because if $L(h)=\langle h,h_0\rangle=\langle h,h_0'\rangle$ then $h_0-h_0^{'}\perp\mathscr{H}$ and so $h_0-h_0'\in\mathscr{H}^{\perp}=\{0\}$. □

3.4. Orthonormal Sets of Vectors and Bases

Definition. (4.1) An **orthonormal** subset of a Hilbert space $\mathscr{H}$ is a subset $\mathscr{E}$having the properties that a) $\|e\|=1\forall e\in\mathscr{E}$ and b) if $e_1\neq e_2$ in $\mathscr{E}$ then $e_1\perp e_2$.

A **basis** for $\mathscr{H}$ is a maximal orthonormal set. (i.e. it is an orthonormal set that is not a subset of any other orthonormal set)

Definition. A **Hamel basis** is a maximal linearly independent set. These are different than orthonormal bases.

Proposition. *(4.2.) If $\mathscr{E}$ is an orthonormal set in $\mathscr{H}$, then there is a basis for $\mathscr{H}$ that contains $\mathscr{E}$.*

Proof. This is an application of Zorn's lemma, just order the orthonormal sets by inclusion. □

Example. (4.3) In $\mathscr{H}=L^2_{\mathbb{C}}[0,2\pi]$, the functions $e_n(t)=(2\pi)^{-1/2}\exp\left(int\right)$ are an orthonormal set. We will see later that these are in fact a basis.

Proposition. *(4.6) (Gram-Schmidt Orthonogonalization Process) If $\mathscr{H}$ is a Hilbert space and $\{h_n:n\in\mathbb{N}\}$ is a lineraly independent subset of $\mathscr{H}$, then there is an orthonormal set $\{e_n:n\in\mathbb{N}\}$ such that for every n, the linear space of $\{e_1,\ldots,e_n\}$equals the linear span of $\{h_1,\ldots,h_n\}$.*

Proof. Its the same proof as the usual Gram-Schmidt process. □

Proposition. *(4.7) Let $\{e_1,\ldots,e_n\}$ be an orthonormal set in $\mathscr{H}$ and let $\mathscr{M}=\overline{span\{e_1,\ldots,e_n\}}$. If P is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$, then:*

$$Ph=\sum_{k=1}^{n}\langle h,e_k\rangle e_k$$

for all $h\in\mathscr{H}$.

Proof. Let $Qh=\sum_{k=1}^{n}\langle h,e_k\rangle e_k$ and check that $h-Qh\perp e_j$ for each j. □

Proposition. *(4.8.) (Bessel's Inequality) If $\{e_n:n\in\mathbb{N}\}$ is an orthonormal set and $h\in\mathscr{H}$ then:*

$$\sum_{n=1}^{\infty}\left|\langle h,e_n\rangle\right|^2\leq\|h\|^2$$

PROOF. For any fixed n, let $h_n = h - \sum_{k=1}^{n} \langle h, e_k \rangle e_k = h - P_n h$. By Pythagoras:

$$\begin{aligned} \|h\|^2 &= \|h_n\|^2 + \left\| \sum_{k=1}^{n} \langle h, e_k \rangle e_k \right\|^2 \\ &= \|h_n\|^2 + \sum_{k=1}^{n} |\langle h, e_k \rangle|^2 \\ &\geq \sum_{k=1}^{n} |\langle h, e_k \rangle|^2 \end{aligned}$$

□

COROLLARY. *(4.9) If $\mathscr{E}$ is an orthonormal set in $\mathscr{H}$ and $h \in \mathscr{H}$ then $\langle h, e \rangle \neq 0$ for at most countably many $e \in \mathscr{E}$*

PROOF. Look at the sets $E_n = \left\{ e \in \mathscr{E} : |\langle e, h \rangle| \geq \frac{1}{n} \right\}$, each is finite by Bessel's inequality. □

COROLLARY. *(4.10) If $\mathscr{E}$ is an orthonormal set (not nessisarily countable) and $h \in \mathscr{H}$ then:*

$$\sum_{e \in \mathscr{E}} |\langle h, e \rangle|^2 \leq \|h\|^2$$

PROOF. Restrict our attention to the $\{e \in \mathscr{E} : \langle h, e \rangle \neq 0\}$ which is countable by the last corrolarty, and now it is just a straight up use of Bessel's ineq. □

REMARK. To make sense of sums over arbitary (possibly uncountable) sets, $\sum_{\alpha \in I}$, order the subsets of I by inclusion, and then treat this as a net. We say the sum converges if this net converges. This will end up being something like absolute convergence.

LEMMA. *(4.12) If $\mathscr{E}$ is an orthonormal set and $h \in \mathscr{H}$ then:*

$$\sum \{\langle h, e \rangle e, e \in \mathscr{E}\}$$

Converges in $\mathscr{H}$

PROOF. Let $e_1, \ldots$ be an enumeration of the elements from $\mathscr{E}$ for which $\langle h, e \rangle \neq 0$. By Bessel's ineq, $\sum |\langle h, e_n \rangle|^2 \leq \|h\|^2 < \infty$.

Now, for any $\epsilon > 0$ take N so large so that $\sum_{i=N}^{\infty} |\langle e_i, h \rangle|^2 \leq \epsilon^2$ and let $F_0 = \{e_1, \ldots, e_N\}$ Then for any $F, G \subset \{\text{all finite subsets of } \mathscr{E}\}$ define $h_F, h_G$by$\sum_{e \in F} \langle h, e \rangle e$. Notice that:

$$\begin{aligned} \|h_F - h_G\|^2 &= \sum \left\{ |\langle h, e \rangle|^2 : e \in F \Delta G \right\} \\ &\leq \sum_{n=N}^{\infty} |\langle h, e_n \rangle|^2 < \epsilon^2 \end{aligned}$$

So this net is a Cauchy net, which means it is convergent. □

THEOREM. *(4.13) Let $\mathscr{E}$ be an orthonormal set in a Hilbert space $\mathscr{H}$. The following are equivalent:*

a) $\mathscr{E}$ is a orthonormal basis.

b) $h \perp \mathscr{E} \implies h = 0$

c)$\overline{span\mathscr{E}} = \mathscr{H}$
d)$h = \sum \{\langle h, e \rangle e : e \in \mathscr{E}\} \forall h \in \mathscr{H}$
e) $\langle g, h \rangle = \sum \{\langle g, e \rangle \langle e, h \rangle : e \in \mathscr{E}\}$
f) $h \in \mathscr{H}$ *then* $\|h\|^2 = \sum \left\{ |\langle h, e \rangle|^2 : e \in \mathscr{E} \right\}$ *(Parseval's Identity)*

PROOF. a) $\implies$ b): If by contradiction there is a non-zero $h \perp \mathscr{E}$, then add this to the set to get a larger orthonormal set.

b) $\iff$ c): We showed that a subsapace is dense if and only if its perpendicular space is trivial. (Cor 2.11) This is exactly that statement!

b) $\implies$ d) : For any h, the vector $h - \sum \{\langle h, e \rangle e : e \in \mathscr{E}\}$ is verified to be in $\mathscr{E}^{\perp}$ and we get the desired result.

d) $\implies$ e): Write g, h as above. Use some care in the definiton of our convergent nets to check that it is indeed true.

e) $\implies$ f) Put $g = h$ to get it!

f) $\implies$ a) Suppose by contradiction that $\mathscr{E}$ is not a basis. Then find an element $e \notin \mathscr{E}$ which is orthonormal to everything. This element will not satisfy Parseval's Identity because the LHS is 1 while the RHS is 0. □

PROPOSITION. *(4.14) If* $\mathscr{H}$ *is a Hilbert space, any two bases have the same cardinality.*

PROOF. If they are finite, then the result is just that from linear algebra. Otherwise, create an injectionby, $e \to \{f \in : \langle e, f \rangle \neq 0\}$, this is a countable set, so this shows that $|\mathscr{E}| \leq |\aleph_0| = ||$. The other direction is the same. □

DEFINITION. (4.15) The cardinality of a basis is called the **dimension** of the Hilbert space.

PROPOSITION. *(4.16) If* $\mathscr{H}$ *is an infintie dimensional Hilbert space, then* $\mathscr{H}$ *is seperable if and only if* $\dim \mathscr{H} = \aleph_0$

PROOF. I'm skipping this! □

REMARK. There is the stuff on Hamel basis's here....that they are uncountable and so on, that I'm skipping.

3.5. Isomorphic Hilbert Spaces and the Fourier Transform

DEFINITION. (5.1) We say that a map $U : \mathscr{H} \to \mathscr{K}$ is an **isomorphism** if it is a linear surjection that preserves the inner product:

$$\langle Ug, Uh \rangle = \langle g, h \rangle$$

PROPOSITION. *(5.2) If* $V : \mathscr{H} \to \mathscr{K}$ *is a linear isometry (i.e.* $\|h - g\| = \|V(h - g)\|$*), then* V *is actually preserves the inner product.*

PROOF. From the polar identity $\|h + \lambda g\|^2 = \|h\|^2 + 2\text{Re}\lambda \langle h, g \rangle + \|g\|^2$ we can get that the inner products actually agree too. □

REMARK. An isometry need not be an isomorphism, because it might not be a surjection. Example: the shift operator from $\ell^2 \to \ell^2$.

THEOREM. *(5.4) Two Hilbert spaces are isomorphic if and only if they have the same dimension*

PROOF. ($\Rightarrow$) If $\mathscr{E}$ is a basis, then it is easy to see that $\{Ue : e \in \mathscr{E}\}$ is a basis too.

($\Leftarrow$) Let $\mathscr{E}$ be any basis for a Hilbert space $\mathscr{H}$. We will show that $\mathscr{H}$ is isomorphic to $\ell^2(\mathscr{E}) = \left\{f : \mathscr{E} \to \mathbb{F} : \sum_{e\in\mathscr{E}} f(e)^2 < \infty\right\}$. For any $h \in \mathscr{H}$, define $\hat{h} : \mathscr{E} \to \mathbb{F}$ by $\hat{h}(e) = \langle h, e\rangle$. By Parseval's identity, $\hat{h} \in \ell^2(\mathscr{E})$ and $\|h\| = \left\|\hat{h}\right\|_{\ell^2}$. The map $U : \mathscr{H} \to \ell^2(\mathscr{E})$ by $Uh = \hat{h}$ is easily verified to be linear, and it is an isometry by the observation we just made $\|h\| = \left\|\hat{h}\right\|_{\ell^2}$. Finally, we see that the range of U is dense in $\ell^2(\mathscr{E})$ because it contains all the indicators δ_e for $e \in \mathscr{E}$.

So indeed, this is an isomorphism. □

COROLLARY. *(5.5) All seperatble infinite dimensional Hilbert spaces are isomorphic.*

3.5.1. Fourier Analysis on the Circle.

REMARK. I did a pretty lazy job with this section.

THEOREM. *(5.6.) If $f : \partial D \to \mathbb{C}$ is a continuous function, then there is a sequence of polynomials $p_n(z, \bar{z})$ so that $p_n(z, \bar{z}) \to f(z)$ uniformoly on ∂D*

REMARK. This can be seen by the Stone-Weirestrass theorem on the algebra of trigonometric functions $\sum_{k=-m}^{m} \alpha_k e^{ik\theta}$.

THEOREM. *(5.7) The set of functions $e_n(t) = (2\pi)^{-1/2} \exp(int)$ is an orthonormal basis for $L^2[0, 2\pi]$.*

PROOF. We will show that the closure (under the uniform norm) of the functions e_n is the whole space $L^2[0, 2\pi]$. This is exactly the last theorem.... □

THEOREM. *This basis gives rise to the map $U : L^2[0, 2\pi] \to \ell^2(\mathbb{Z})$ by $U : f \to \hat{f}$ with $\hat{f} = \langle f, e_n\rangle = \int f(t)e^{-int}$. This map is a linear isometry.*

3.6. Direct Sum of Hilbert Spaces

This section just tells you how to define an inner product on the direct sum of Hilbert spaces,

$$\langle h_1 \oplus k_1, h_2 \oplus k_2\rangle := \langle h_1, h_2\rangle + \langle k_1, k_2\rangle$$

The main thing to be done here is to extend this to infinite sums:

PROPOSITION. *(6.2) If $\mathscr{H}_1, \ldots$ are Hilbert spaces, let $\mathscr{H} = \left\{(h_n) : h_n \in \mathscr{H}_n \forall n \ \sum \|h_n\|^2 < \infty\right\}$, then the inner product:*

$$\langle h, g\rangle_{\mathscr{H}} := \sum_{n=1}^{\infty} \langle h_n, g_n\rangle_{\mathscr{H}_n}$$

This inner product makes $\mathscr{H}$ a Hilbert space.

This Hilbert space is denoted $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2 \oplus \ldots$

Operators on Hilbert Spaces

These are notes from Chapter 2 of [**1**].

4.7. Basic Stuff

PROPOSITION. *(1.1) Let $\mathscr{H}$ and $\mathscr{K}$ be Hilbert spaces and $A : \mathscr{H} \to \mathscr{K}$ a linear transformation. The following are equivalent:*
a) A is continuous
b) A is continuous at 0
c) A is continuous at some point
d) There is a constant $c > 0$ such that $\|Ah\| \leq c\,\|h\|$ for all h

PROOF. Similar to the proof for functionals we did earlier. □

DEFINITION. An operator $A : \mathscr{H} \to \mathscr{K}$ is called bounded if$\|A\| := \sup_{\|h\|=1} \|Ah\| < \infty$. The space of all bounded operators $A : \mathscr{H} \to \mathscr{K}$ is denoted $\mathscr{B}(\mathscr{H}, \mathscr{K})$.

PROPOSITION. *(1.2) a) $\|A + B\| \leq \|A\| + \|B\|$*
b) $\|\alpha A\| = |\alpha|\,\|A\|$
c) $\|AB\| \leq \|A\|\,\|B\|$

PROOF. Follows your nose from the definition. □

PROPOSITION. *(Schur Test) On $\ell^2(\mathbb{N})$, let $\alpha_{ij} := \langle Ae_i, e_j \rangle$. Suppose that $\exists p_i > 0$ and β,$\gamma > 0$ with:*

$$\begin{aligned} \sum_i \alpha_{ij} p_i &\leq \beta p_j \\ \sum_j \alpha_{ij} p_j &\leq \gamma p_i \end{aligned}$$

Then $\|A\|^2 \leq \beta\gamma$

PROOF. Still working on this one! □

THEOREM. *(1.5.) Let (X, Ω, μ) be a $\sigma-$finite measure space and put $\mathscr{H} = L^2(X, \Omega, \mu)$ For $\phi \in L^\infty(\mu)$ define $M_\phi : L^2(\mu) \to L^2(\mu)$ by $M_\phi f = \phi f$. Then $M_\phi \in \mathscr{B}(L^2(\mu))$ and $\|M_\phi\| = \|\phi\|_\infty$ where $\|\cdot\|_\infty$ is the essential supremum norm with respect to the measure μ.*

PROOF. The fact that $\|M_\phi\| \leq \|\phi\|_\infty$ is clear since $|\phi| \leq \|\phi\|_\infty$a.e. . On the other hand for any $\epsilon > 0$, we can find a positive measure set so that $|\phi| > \|\phi\|_\infty - \epsilon$ and then take some L^2 functions concentrated here to get the other inequality. □

THEOREM. *(1.6.) let (X,Ω,μ) be a positive measure space and suppose $k : X\times X\to\mathbb{F}$ is a $\Omega\times\Omega$ measurable function for which there are constants c_1 and c_2 so that :*

$$\begin{aligned}\int |k(x,y)|\, d\mu(y) &\leq c_1 \text{ for a.e. } x\\ \int |k(x,y)|\, d\mu(x) &\leq c_2 \text{ for a.e. } y\end{aligned}$$

Then define $K : L^2(\mu)\to L^2(\mu)$ by:

$$(Kf)(x) = \int k(x,y)f(y)d\mu(y)$$

Then K is a bounded linear operator with $\|K\|\leq (c_1c_2)^{1/2}$

PROOF. The trick is to use Cauchy Schwaz:

$$\begin{aligned}|Kf(x)| &\leq \int |k(x,y)|\,|f(y)|\,\mathrm{d}\mu(y)\\ &= \int |k(x,y)|^{1/2}\,|k(x,y)|^{1/2}\,|f(y)|\,\mathrm{d}\mu(y)\\ &\leq \left[\int |k(x,y)|\,\mathrm{d}\mu(y)\right]^{1/2}\left[\int |k(x,y)|\,|f(y)|^2\,\mathrm{d}\mu(y)\right]^{1/2}\\ &\leq c_1^{1/2}\left[\int |k(x,y)|\,|f(y)|^2\,\mathrm{d}\mu(y)\right]^{1/2}\end{aligned}$$

And so now integrating over x now gives:

$$\begin{aligned}\int |Kf(x)|^2\,\mathrm{d}\mu(x) &\leq c_1\int\int |k(x,y)|\,|f(y)|^2\,\mathrm{d}\mu(y)\mathrm{d}\mu(x)\\ &= c_1\int |f(y)|^2\left(\int |k(x,y)|\,\mathrm{d}\mu(x)\right)\mathrm{d}\mu(y) \text{ by Fubini-Tonnelli}\\ &\leq c_1c_2\,\|f\|^2\end{aligned}$$

□

4.8. Adjoint of an Operator

REMARK. I made the executive descision to swithc from $\mathscr{H}$ to $\mathcal{H}$ as the symbol to be used for a Hilbert space.

DEFINITION. (2.1) We say that $u : \mathcal{H}\times\mathcal{K}\to\mathbb{F}$ is sesquilinear if it is bilinear except for conjugation in the second component.

EXAMPLE. For any bounded operator A, the form $u(x,y) = \langle Ax, y\rangle$ is a sesquilinear form. This can be shown by the properties of the inner product.

THEOREM. *(2.2) If $u : \mathcal{H}\times\mathcal{K}\to\mathbb{F}$ is a bounded sesquilinear form with bound $|u(h,k)|\leq M\,\|h\|\,\|k\|$ then there are unique operators $A\in\mathscr{B}(\mathcal{H},\mathcal{K})$, $B\in\mathscr{B}(\mathcal{K},\mathcal{H})$ so that:*

$$u(h,k) = \langle Ah, k\rangle = \langle h, Bk\rangle$$

and $\|A\|, \|B\|\leq M$

PROOF. The idea is to use Riesz. For fixed h, check that $\overline{u(h,\cdot)}$ is a linear functional on $\mathcal{K}$ (holds since u is given to be sesquilinear) Hence, by the Riesz representation theorem, there is an element k (depending on h) so that $\overline{u(h,\cdot)} = \langle \cdot, k\rangle$. Check by using the uniqueness and linearity that the map $A : h \to k$ is a linear map, and it is boudned because u is bounded. The same stuff works to show B exists. □

DEFINITION. (2.4.) For a given $A \in \mathscr{B}(\mathcal{H},\mathcal{K})$, we can always find a $B \in \mathscr{B}(\mathcal{K},\mathcal{H})$ with $\langle Ah, k\rangle = \langle h, Bk\rangle$. This matrix B is called the **adjoint** of A and is often denoted A^*.

PROPOSITION. *If $U \in \mathscr{B}(\mathcal{H},\mathcal{K})$ then U is an isomorphism if and only if U is invertible and $U^{-1} = U^*$*

PROOF. Suppose U is invertible. We have only to verify then that $\langle Uf, Ug\rangle = \langle f, g\rangle$ if and only if $U^{-1} = U^*$. Indeed, it is always true that $\langle Uf, Ug\rangle = \langle f, U^*Ug\rangle$ and then:

$$\begin{aligned}\langle f, U^*Ug\rangle = \langle f, g\rangle &\iff \langle f, (Id - U^*U)g\rangle = 0\\ &\iff Range(Id - U^*U) \subset \mathcal{H}^\perp = \{0\}\\ &\iff U^*U = Id\end{aligned}$$

□

PROPOSITION. *If $A, B \in \mathscr{B}(\mathcal{H})$ and $\alpha \in \mathbb{F}$ then:*
a) $(aA + B)^* = \bar{a}A^* + B^*$
b) $(AB)^* = B^*A^*$
c) $(A^*)^* = A$
d) If A is invertble in $\mathscr{B}(\mathcal{H})$ then A^* is invertble with:

$$(A^*)^{-1} = \left(A^{-1}\right)^*$$

PROOF. Pretty standard exercise! □

PROPOSITION. *If $A \in \mathscr{B}(\mathcal{H})$ then $\|A\| = \|A^*\| = \|A^*A\|^{1/2}$*

PROOF. Have for any h with $\|h\| = 1$ that:

$$\begin{aligned}\|Ah\|^2 &= \langle Ah, Ah\rangle\\ &= \langle A^*Ah, h\rangle\\ &\le \|A^*Ah\|\,\|h\|\\ &\le \|A^*A\|\,\|h\|\,\|h\|\\ &\le \|A^*\|\,\|A\| \cdot 1 \cdot 1\end{aligned}$$

Hence $\|A\|^2 \le \|A^*A\| \le \|A^*\|\,\|A\|$. Canceling $\|A\|$, we have $\|A\| \le \|A^*\|$. The argument holds equally well the other way (or consdier that $A^{**} = A$) and so we have $\|A^*\| \le \|A\|$. Hence we must actually have equality everywhere and this proves the claim. □

EXAMPLE. Conway shows a few nice examples here, including the forward shift on $\ell^2(\mathbb{N})$ whose adjoint is the backwards shift.

DEFINITION. (2.11) If $A \in \mathscr{B}(\mathcal{H})$ we say that A is **hermitian** or **self-adjoint** if $A^* = A$ and we say that A is **normal** if $AA^* = A^*A$

REMARK. If we think of *as being analogous to conjugation on the complex numbers, then Hermitain operators are the analogous of the real numebrs. Normal operators are the true analogous of arbitarty complex numbers, the analogy doesnt really make sense for non-normal operators.

PROPOSITION. *(2.12) $A \in \mathscr{B}(\mathcal{H})$ is Hermitian if and only if $\langle Ah, h\rangle \in \mathbb{R}$ for all $h \in \mathcal{H}$ (This only works for $\mathbb{C}$ valued Hilbert spaces)*

PROOF. ($\Rightarrow$) $\langle Ah, h\rangle = \langle h, A^*h\rangle = \langle h, Ah\rangle = \overline{\langle Ah, h\rangle}$
($\Leftarrow$)Assume $\langle Af, f\rangle \in \mathbb{R}$ for all f For $h, g \in \mathcal{H}$ and $\alpha \in \mathbb{C}$ consider:

$$\langle A(h+\alpha g), h+\alpha g\rangle = \langle Ah, h\rangle + \alpha\langle Ah, g\rangle + \overline{\alpha}\langle Ag, h\rangle + |\alpha|^2\langle Ag, g\rangle$$

The LHS and two terms on the RHS are $\mathbb{R}$ by hypothesis. Hence $\alpha\langle Ah, g\rangle + \overline{\alpha}\langle Ag, h\rangle$ is real too, so it is equal to its complex conjugate. On the other hand:

$$\begin{aligned}\alpha\langle Ah, g\rangle + \overline{\alpha}\langle Ag, h\rangle &= \bar{\alpha}\langle g, Ah\rangle + \alpha\langle h, Ag\rangle \\ &= \bar{\alpha}\langle A^*g, h\rangle + \alpha\langle A^*h, g\rangle\end{aligned}$$

If you put in $\alpha = 1$ first and then $\alpha = i$, you get two linear equations and two unknowns that leads us to $\langle Ag, h\rangle = \langle A^*g, h\rangle$. Since this holds for every g, h it must be that $A = A^*$ □

PROPOSITION. *(2.13) If $A = A^*$ then:*

$$\|A\| = \sup_{\|h\|=1} |\langle Ah, h\rangle|$$

PROOF. (my idea: The idea is to use the fact which is always true that $\|A\| = \sup_{\|h\|=1,\|g\|=1} |\langle Ah, g\rangle|$. (This comes from $\|x\| = \sup_{\|y\|=1} |\langle x, y\rangle|$). From this it is clear that $\|A\| \geq \sup_{\|h\|=1} |\langle Ah, h\rangle|$. To see the other inequality, do a change of variable so that $x = \frac{g+h}{2}$, $y = \frac{g-h}{2}$. (The main idea is to manipulate $\langle Ah, g\rangle$ into $\langle Ax, x\rangle + \langle Ay, y\rangle + \langle Ax, y\rangle - \langle Ay, x\rangle$, then use the Hermitain-ness to see the two cross terms as conjugate conjugates, $\langle Ax, y\rangle = \overline{\langle Ay, x\rangle}$ since $A = A^*$.) By the parralelogram law, $2\|x\|^2 + 2\|y\|^2 = \|h\|^2 + \|g\|^2 = 2$. So $\|x\|^2 = r^2$ and $\|y\|^2 = 1 - r^2$ for some $0 \leq r \leq 1$. So we have:

$$\|A\| = \sup_{0<r<1} \sup_{\|x\|=r, \|y\|=\sqrt{1-r^2}} |\langle Ax, x\rangle + \langle Ay, y\rangle + 2i\mathrm{Im}\langle Ax, y\rangle|$$

Now, $\langle Ax, x\rangle$,$\langle Ay, y\rangle$ are real while the other term is imaginary, so they split up nicely. The fact that $\langle Ax, x\rangle$ and $\langle Ay, y\rangle$ are real is particularly nice! By scaling, when $\|x\| = r$ we have $\sup_{\|x\|=r}\langle Ax, x\rangle = r^2\sup_{\|z\|=1}\langle Az, z\rangle$, which controls the real part. ......) □

The way Conway does it is to use $A = A^*$to get to:

$$4\mathrm{Re}\langle Ah, g\rangle = \langle A(h+g), h+g\rangle - \langle A(h-g), h-g\rangle$$

So if $M = \sup_{\|h\|=1}\langle Ah, h\rangle$ then have by scaling:

$$\begin{aligned}4\mathrm{Re}\langle Ah, g\rangle &\leq M\|h+g\|^2 + M\|h-g\|^2 \\ &= 2M\|h\|^2 + 2M\|g\|^2 \text{ by parralelogram law} \\ &= 4M\end{aligned}$$

By rotating g or h approproatly, this argument can be modified from the concluison $\mathrm{Re}\langle Ah, g\rangle \leq M$ to $|\langle Ah, g\rangle| \leq M$.

COROLLARY. *(2.14) If $A = A^*$ and $\langle Ah, h\rangle = 0$ for every h, then $A = 0$.*

REMARK. In a complex Hilbert space, $\langle Ah, h\rangle \in \mathbb{R}$ for every h implies $A = A^*$, so this condition could be dropped from this corollary in the complex setting.

PROPOSITION. *(2.16) If $A \in \mathscr{B}(\mathcal{H})$ the following are equivalent:*
a) A is normal
b) $\|Ah\| = \|A^*h\|$ for all h
If we are working in $\mathbb{C}$ then this is also equivalent to:
c) Th real and imaginary parts of A commute

PROOF. Notice that:

$$\begin{aligned} \|Ah\|^2 - \|A^*h\|^2 &= \langle Ah, Ah\rangle - \langle A^*h, A^*h\rangle \\ &= \langle (A^*A - AA^*)h, h\rangle \end{aligned}$$

So the equivalence of a) and b) follows from the previous corollary.

To check the last bit, let $B = \mathrm{Re}A$ and $C = \mathrm{Im}A$ so that $A = B + iC, A^* = B - iC$ and then write out:

$$\begin{aligned} A^*A &= B^2 - iCB + iBC + C^2 \\ AA^* &= B^2 + iCB - BC + C^2 \end{aligned}$$

So that the two are equal if and only if $BC = CB$. □

PROPOSITION. *(2.17) The following are equivalent:*
a) A is an isometry
b) $A^*A = I$
c) $\langle Ah, Ag\rangle = \langle h, g\rangle$ for all $h, g \in \mathcal{H}$

PROOF. a) $\iff$ c) was discussed earlier (basically because the inner product can be written purely in terms of norms via the parallelogram law), and b) $\iff$ c) is clear because $\langle Ah, Ag\rangle = \langle h, g\rangle \iff \langle (A^*A - Id)h, g\rangle = 0$. □

PROPOSITION. *(2.18) The following are equivalent:*
a) $A^*A = AA^* = I$
b) A is unitary (That is: A is a surjective isometry)
c) A is a normal isometry

PROOF. a) $\implies$ b): From the hypothesis a), we know that A is invertable and from the previous propsition it is an isometry. Hence it is a surjective isometry.

b) $\implies$ c) By Prop 2.17, $A^*A = I$. When A is a surjective isometry, the inverse of A must also be a surjective isometry, so we have also from Prop 2.17 that $I = \left(A^{-1}\right)^* A^{-1}$. Now use $(A^*)^{-1} = \left(A^{-1}\right)^*$ to manipulate this into $I = (AA^*)^{-1}$. So then $AA^* = A^*A = I$ and A is normal.

c) $\implies$ a) By Prop 2.17, since A is an isometry, $A^*A = I$, since A is also normal we have $A^*A = AA^* = I$ as desired. □

THEOREM. *(2.19) If $A \in \mathscr{B}(\mathcal{H})$ then $\ker A = (ranA^*)^\perp$*

PROOF. If $h \in \ker A$ and $g \in \mathcal{H}$ then $\langle h, A^*g\rangle = \langle Ah, g\rangle = \langle 0, g\rangle = 0$. This shows $\ker A \subset (\mathrm{ran}A^*)^\perp$

If $h \perp \mathrm{ran}A^*$ then $\langle h, A^*g\rangle = 0$ for every $g \in \mathcal{H}$ and so $\langle Ah, g\rangle = 0$ for every g, and hence $Ah \in \mathcal{H}^\perp = \{0\}$. This shows ran $A^{*\perp} \subset \ker A$. □

4.9. Projections and Idempotents; Invariant and Reducing Subspaces

Definition. (3.1) An **idempotent** is a bounded linear operator E so that $E^2 = E$. A **orthogonal projection** is an idempotent P such that $\ker P = (\mathrm{ran} P)^{\perp}$. We actually use the word **projection** to refer only to orthogonal projections.

Example. An *non-orthogonal* projection is an idempotent, but it is not a proejction. For example, take an basis (not nessisarily orthonormal) for $\mathbb{R}^n$, say $e_1, \ldots, e_n$ then every vector v has a unique writting $v = \sum \alpha_k e_k$, and the map $P : v \to \alpha_i e_i$ for some fixed i is a "projection" but it is not an orthogonal projection unless the basis is an orthogonal one. (In this case $\ker P = \mathrm{span}\left\{e_1, \ldots, \hat{e}_i, \ldots, e_n\right\}$ while $(\mathrm{ran} P)^{\perp} = \mathrm{span}\left\{e_i\right\}^{\perp}$)

One thing to notice in this case is that $\|P\| > 1$. Take any vector $v = \sum \alpha_k e_k$ so that $\|v\|^2 < \alpha_k \|e_k\|$ (indeed the existence of this follows by the "cosine law" in a Hilbert space), and the we will have that $\|Pv\| = \alpha_k$

Proposition. *a) E is an idempotent $\iff I - E$ is an idempotent*

b) If E is an idempotent, $ran(E) = \ker(I - E)$ and $\ker E = ran(I - E)$ and $ran(E)$ is a closed linear subspaces of $\mathcal{H}$

c) If $\mathcal{M} = ranE$ and $\mathcal{N} = \ker E$ then $\mathcal{M} \cap \mathcal{N} = \{0\}$ and $M + N = \mathcal{H}$

Proof. Check that $I - E$ is idempotent by verifying that $(I - E)^2 = \ldots = I - E$. Then $\mathrm{ran}(E) = \ker(I - E)$ since $(I - E)h = 0 \iff h = Eh \iff h \in \mathrm{ran} E$ (Notice $h \in \mathrm{ran} E \implies h = Eg \implies Eh = E^2 g = Eg = h$) . Since $\ker(A)$ is a closed linear subspace for any operator A, $\mathrm{ran}(E) = \ker(I - E)$ shows that $\mathrm{ran}(E)$ is a closed linear subspace too. □

Remark. In general the range of an operator is NOT a closed subapce. The Kernal of an operator always is. For example the diagonal operator from on $\ell^2(\mathbb{N})$ by $Ae_n = \frac{1}{n} e_n$. (The range is dense in ℓ^2 but contins only sequences for which $\sum n \langle x, e_n \rangle < \infty$ for example $\left(1, \frac{1}{2}, \ldots\right) \notin \mathrm{ran} A$). Notice that this operator has a point in its spectrum that is not an eigenvalue....maybe this is relevant?

Lemma. *(Mini Lemma) If E is idempotent then:*

$$h \in ranE \iff h = Eh$$

Proof. ($\Leftarrow$) is clear. ($\Rightarrow$)If $h = Eg$ for some g, then apply E to both sides to get $Eh = E^2 g = Eg$ but $Eg = h$ so this says $Eh = Eg = h$. □

Proposition. *(3.3.) Suppose E is an* **idempotent** *on $\mathcal{H}$ and $E \neq 0$, the following are equivalent:*

a) E is a projection (i.e. $\ker E = (ranE)^{\perp}$ is the definition)

b) E is the orthogonal projection of $\mathcal{H}$ onto ranE

c) $\|E\| = 1$

d) E is hermitian, $E = E^*$

e) E is normal, $EE^* = E^* E$

f) $\langle Eh, h \rangle \geq 0$ fora all $h \in \mathcal{H}$

Proof. a) $\implies$ b): Have $h - Eh = (I - E)h \in \ker E = (\mathrm{ran} E)^{\perp}$. By the uniqueness of the orhogonal projection Eh is the orthogonal projection on $\mathrm{ran} E$. (Recall: a projection is the unique operator so that $h - P_{\mathcal{M}} h \in \mathcal{M}^{\perp}$ for all h...see thm 2.7)

b) $\Longrightarrow$ c): Follows from the fact that orthogonal projections have $\|P_{\mathcal{M}}h\| \leq \|h\|$ with equality for $h \in \mathcal{M}$.

c) $\Longrightarrow$ a): Take any $h \in (\ker E)^{\perp}$, we know $h - Eh \in \ker E$ so we have $\langle h, h - Eh \rangle = 0 \implies \|h\|^2 = \langle Eh, h\rangle$ for any $h \in (\ker E)^{\perp}$. On the other hand, by C.S. we know $|\langle Eh, h\rangle| \leq \|Eh\| \|h\| \leq \|E\| \|h\|^2 = \|h\|^2$ here, so we have an equality sandwhich and we conclude that for any $h \in (\ker E)^{\perp}$that $\|Eh\| = \|h\| = \langle Eh, h\rangle^{1/2}$.

Now by the polarization identity, for any $h \in \ker E^{\perp}$have:

$$\|h - Eh\|^2 = \|h\|^2 - 2\text{Re}\,\langle Eh, h\rangle + \|Eh\|^2 = 0$$

Hence $h \in (\ker E)^{\perp} \implies h = Eh \iff h \in \text{ran}E$. This shows $\ker E^{\perp} \subset \text{ran}E$.

Conversely, if $g \in \text{ran}E$ then write $g = P_{\ker E}g + P_{\ker E^{\perp}}g$. Since $g = Eg$, and $E(P_{\ker E}g) = 0$, have then $g = Eg = 0 + E\left(P_{\ker E^{\perp}}g\right)$. Since $P_{\ker E^{\perp}}g \in \ker E^{\perp} \subset \text{ran}E$ by the abover argument, we know $E(P_{\ker E^{\perp}}g) = P_{\ker E^{\perp}}g$, so we have $g = E(P_{\ker E^{\perp}}g) = P_{\ker E^{\perp}}g \in \ker_{E^{\perp}}$.

.....

I'm going to skip the rest of this proof and this section for now

....

The next bit basically has two definitions and a small result about them: □

Definition. (3.5) Given a closed subspace $\mathcal{M}$ and its orthogonal space $\mathcal{M}^{\perp}$ any operator A can be decomped as $A = IAI = (P_{\mathcal{M}} + P_{\mathcal{M}^{\perp}})\,A\,(P_{\mathcal{M}} + P_{\mathcal{M}^{\perp}}) = P_{\mathcal{M}}AP_{\mathcal{M}} + P_{\mathcal{M}}AP_{\mathcal{M}^{\perp}} + P_{\mathcal{M}^{\perp}}AP_{\mathcal{M}} + P_{\mathcal{M}^{\perp}}AP_{\mathcal{M}^{\perp}} := W + X + Y + Z$. In matrix form this is $A = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$ where the first row/column represents $\mathcal{M}$ and the second represents $\mathcal{M}^{\perp}$

Definition. (3.4. and Prop 3.7) We say that $\mathcal{M}$ is **invariant** for A if $A\mathcal{M} \subset \mathcal{M}$. The following are equivalent:

a) $\mathcal{M}$ is invariant for A (i.e. $A\mathcal{M} \subset \mathcal{M}$)
b) $P_{\mathcal{M}}AP_{\mathcal{M}} = AP_{\mathcal{M}}$
c) $Y = 0$ in the above

Definition. (3.4 and Prop 3.7) We say that $\mathcal{M}$ is **reducing** for A if $A\mathcal{M} \subset \mathcal{M}$ and $A\mathcal{M}^{\perp} \subset \mathcal{M}^{\perp}$. The following are equivalent:

a) $\mathcal{M}$ reduces A (i.e. $A\mathcal{M} \subset \mathcal{M}$ and $A\mathcal{M}^{\perp} \subset \mathcal{M}^{\perp}$)
b) $P_{\mathcal{M}}A = AP_{\mathcal{M}}$
c) X and Y are both 0 from the above
d) $\mathcal{M}$ is invariant for both A and A^*

Proof. Not too hard...just follow your nose mostly! □

4.10. Compact Operators

Symbol	Name	Definition
$\mathscr{B}(\mathcal{H}, \mathcal{K})$	Bounded Operators	$\|T\| < \infty$i.e. $T\,(\text{ball } \mathcal{H})$ is <u>bounded</u>
$\mathscr{B}_0(\mathcal{H}, \mathcal{K})$	Compact Operators	$T\,(\text{ball } \mathcal{H})$ is <u>pre-compact</u> (compact closure or "totally bounded")
$\mathscr{B}_{00}\,(\mathcal{H}, \mathcal{K})$	Finite Rank Operators	ran(T) is finite dimensional

Proposition. *(4.2) a)*$\mathscr{B}_0 \subset \mathscr{B}$
b) If $T_n \in \mathscr{B}_0$*,* $T \in \mathscr{B}$ *and* $\|T_n - T\| \to 0$ *then* $T \in \mathscr{B}_0$

c) If $A \in \mathscr{B}(\mathcal{H})$ and $B \in \mathscr{B}(\mathcal{K})$, then for $T \in \mathscr{B}_0(\mathcal{H},\mathcal{K})$ we have $TA, BT \in \mathscr{B}_0(\mathcal{H},\mathcal{K})$

Proof. a) is clear because precompact sets (totally bounded sets) are always bounded

b) We verify directly that $T(\text{ball } \mathcal{H})$ is totally bounded. For $\epsilon > 0$ choose n so large so that $\|T_n - T\| < \epsilon/3$. Since $T_n(\text{ball } \mathcal{H})$ is totally bounded there are vectors $h_1, \ldots, h_m$ so that $T_n h_1, \ldots T_n h_m$ form an $\epsilon/3$ net for $T_n(\text{ball}\mathcal{H})$. Hence for any h with $\|h\| \leq 1$, there is an h_i with $\|Th - Th_i\| < \frac{\epsilon}{3}$. Claim now that $Th_1, \ldots, Th_n$ form an ϵ net for $T(\text{ball}\mathcal{H})$ since:

$$\begin{aligned} \|Th - Th_i\| &\leq \|Th_i - T_n h_i\| + \|T_n h_i - T_n h\| + \|T_n h - Th\| \\ &\leq 2\|T - T_n\| + \epsilon/3 \\ &< \epsilon \end{aligned}$$

c) To see that $TA \in \mathscr{B}_0$ consider as follows. Since A is a bounded operator, can find a ball so that $A(\text{ball}\mathcal{H}) \subset$ bigger ball $\mathcal{H}$, but the $TA(\text{ball}\mathcal{H}) \subset T(\text{bigger ball})$ is a subset of totally bounded set, and is hence totally bounded.

To see that $BT \in \mathscr{B}_0$ we notice that for any totally boudned set K, $B(K)$is totally bounded: get an ϵ net for $B(K)$ by taking the image of an $\epsilon/\|K\|$ net for K through the map B. Hence $B(T(\text{ball}\mathcal{H}))$ is totally boudned by virtue of the fact that $T(\text{ball}\mathcal{H})$ is totally bounded. □

Theorem. *(4.4.) The following are equivalent:*

a) $T \in \mathscr{B}(\mathcal{H},\mathcal{K})$ is compact

b) T^* is compact

*c) There is a sequence $T_n \in \mathscr{B}_{00}$ of operators of **finite rank** so that $\|T_n - T\| \to 0$*

Proof. c) $\implies$ a) is clear since every finite rank operator is compact (In symbols:$\mathscr{B}_{00} \subset \mathscr{B}_0$) and by the previous proposition, a limit (in the norm topology) of compact operators is compact.

a) $\implies$ c) Since $T(\text{ball}\mathcal{H})$ is pre-compact, it is separable (take a sequence of $\frac{1}{n}$-nets). Therefore, $\overline{\text{ran}(T)} =: \mathcal{L}$ is a seperable subspace of $\mathcal{K}$. Take $\{e_1, \ldots\}$ a basis for $\mathcal{L}$ and let P_n be the orthogonal projection onto $\text{span}\{e_k : 1 \leq k \leq n\}$. Put $T_n = P_n T$ and note that each $T_n$has finite rank. To see that $\|T_n - T\| \to 0$ consider as follows.

<u>Claim:</u> If $h \in \mathcal{H}$, then $\|T_n h - Th\| \to 0$

<u>Pf:</u> By definition of $\mathcal{L}$, $Th =: k \in \mathcal{L}$. Now, $\|T_n h - Th\| = \|P_n k - k\| = \sum_{k>n} |\langle e_k, k\rangle|^2 \to 0$ by Parseval identity

Now, we will use the precompactness of $T(\text{ball}\mathcal{H})$ to show that claim is enough to have $\|T_n - T\| \to 0$. Indeed, for any $\epsilon > 0$ take an $\epsilon/3$ net of $T(\text{ball}\mathcal{H})$ call it $Th_1, \ldots, Th_m$. Take n_0 so large so that $\|T_n h_i - Th\| < \epsilon/3$ for every Th_i in the chosen net (this follows by the lemma since there are finitely many of them to deal with). Then for $n \geq n_0$ and any $\|h\| \leq 1$ we find the h_j so that $\|Th - Th_j\| < \epsilon/3$ and we have the following estimate:

$$\begin{aligned} \|Th - T_n h\| &\leq \|Th - Th_j\| + \|Th_j - T_n h_j\| + \|P_n(Th_j - Th)\| \\ &\leq 2\|Th - Th_j\| + \epsilon/3 \text{ by choice of } n_0 \text{ and since } P \text{ a projection} \\ &\leq 2\epsilon/3 + \epsilon/3 \text{ since } Th_j \text{ an } \epsilon/3 \text{ net} \\ &= \epsilon \end{aligned}$$

This estimate holds for every $\|h\| \leq 1$, so we conclude that $\|T - T_n\| \leq \epsilon$ for $n \geq n_0$, and $\|T - T_n\| \to 0$ as desired.

c) $\implies$ b) For $T_n \in \mathscr{B}_{00}$, it is easily verified that $T_n^* \in \mathscr{B}_{00} \subset \mathscr{B}_0$ too and $\|T - T_n\| = \|T^* - T_n^*\| \to 0$ so we see that $T^*$is the limit (norm topology) of compact operators and is hence a compact operator.

b) $\implies$ a) Can do this the sneaky way: b) is the same as a) for T^* so by a) $\implies$ c), we have c) for the operator T^* and then by c) $\implies$ b), we have b) for the operator T^*, which is really a) for the operator T as desired. □

REMARK. There is a slightly less roundabout proof of the fact that T compact $\implies T^*$compact using the Bolazanno-Weirestrass characterization of sequences and the Arzela-Ascoli theorem (By Riesz, $\mathcal{K} \equiv \mathcal{K}^*$so these are really functions)

COROLLARY. *(4.5) If $T \in \mathscr{B}_0$ then $\overline{ranT}$ is separable and if e_n is a basis for $\overline{ranT}$ and P_n is the projection onto the first n basis elements, then $\|P_nT - T\| \to 0$*

PROPOSITION. *(4.6) Let $\mathcal{H}$ be a separable Hilbert space with basis e_n. Suppose that α_n is a sequence with $M := \sup|\alpha_n| < \infty$. If A is the diagonal operator, $Ae_n = \alpha_n e_n$ for all n, then A is a bounded linear operator with $\|A\| \leq M$. (This part is easy, the real proposition is the next bit)*

The operator A is compact if and only if $\alpha_n \to 0$ as $n \to \infty$

PROOF. Take P_n to be the projection onto $e_1, \ldots, e_n$. Then $A_n = A - AP_n$ is diagonalizable with $Ae_j = \alpha_j e_j$ for $j > n$ and $Ae_j = 0$ for $j \leq n$. Now, since $AP_n \in \mathscr{B}_{00}(\mathcal{H})$ and $\|A_n\| = \sup\{|\alpha_j| : j > n\}$ then we know that $\|A_n\| \to 0$ if and only if $\alpha_n \to 0$.

If $\alpha_n \to 0$ we have that $A - A_n$ are all finite rank and $\|(A - A_n) - A\| = \|A_n\| \to 0$ shows A is the limit (norm-top) of finite rank matrices, and is hence compact.

Conversly, if A is compact, then the corollary shows that $\|A - P_nA\| \to 0$ so $\|A_n\| = \|A - P_nA\| \to 0$ and consequently $\alpha_n \to 0$ □

PROPOSITION. *(4.7) For $k \in L^2(X \times X, \Omega \times \Omega, \mu \times \mu)$, the integral operator K defined by:*

$$(Kf)(x) = \int k(x,y)f(y)d\mu(y)$$

is a compact operator and $\|K\| \leq \|k\|_{L^2}$

The proof uses the following lemma:

LEMMA. *(4.8.) If e_i is a basis for $L^2(X, \Omega, \mu)$ and:*

$$\phi_{ij}(x,y) = e_j(x)\overline{e_i(y)}$$

Then ϕ_{ij} is an orthonormal set in $L^2(X\times X, \Omega\times\Omega, \mu\times\mu)$ and $\langle k, \phi_{ij}\rangle_{L^2(X\times X)} = \langle Ke_j, e_i\rangle_{L^2(X)}$

PROOF. (of Prop 4.7) By C.S. it is easy to check that $\|Kf\|^2 \leq \|k\|^2 \|f\|^2$ and so K is bounded.

To do the compactness you let $K_n = KP_n + P_nK - P_nKP_n$ where P_n is the orthogonal projection onto the first n basis elements, and then check that this is finite rank operator. Then do a bit of work to show $\|K_n - K\| \to 0$ (it will be bounded above by the tail of $\langle Ke_j, e_i\rangle$ which is equal to the tail of $\langle k, \phi_{ij}\rangle$ and $\to 0$ by Parseval. □

4.10.1. First look at eigenvalues for bounded operators on a Hilbert space.

DEFINITION. (4.9) If $A \in \mathscr{B}(\mathcal{H})$, a scalar α is called an **eigenvalue** of A if $\ker(A - \alpha Id) \neq \{0\}$. If h is a non-zero vector in $\ker(A - \alpha)$ then h is called an **eigenvector**. The set of eigenvalues for A is denoted by $\sigma_p(A)$

PROPOSITION. *(4.13) If $T \in \mathscr{B}_0$ is a compact operator, and $\lambda \in \sigma_p(T)$ is an eigenvalue and $\lambda \neq 0$, then the eigenspace $\ker(T - \lambda Id)$ is finite dimensional.*

PROOF. Suppose by contradiction that $\ker(T - \lambda Id)$ has an infinite orthonormal sequence e_n. Then, since T is compact, and e_n is a bounded sequence, there is a subsequence Te_{n_k} which converges. But $\left\|Te_{n_k} - Te_{n_j}\right\|^2 = \lambda \left\|e_{n_k} - e_{n_j}\right\|^2 = 2\left|\lambda\right|^2 > 0$ cannot possibly be convergent! This contradiction shows that $\ker(T - \lambda Id)$ is finite dimensional. ☐

PROPOSITION. *(4.14) If T is a compact operator on $\mathcal{H}$ and $\lambda \neq 0$ and $\inf\left\{\left\|(T - \lambda Id)h\right\| : \|h\| = 1\right\} = 0$ then $\lambda \in \sigma_p(T)$*

REMARK. Later on in the book, we will give this type of thing a name. The **approximate point spectrum** is the set where there is a sequence of unit vectors with $\lim_{n\to\infty}\|Tx_n - \lambda x_n\| = 0$ and we denote by σ_{ap} the set of all such λ . In this language, this result says that compact operators don't have anything in $\sigma_{ap} - \sigma_p$. In other words, for compact operators, every approximate eigenvalue is a true eignevalue.

PROOF. If $\|Tx_n - \lambda x_n\| \to 0$, then by the compactness of T there is a convergent subsequence $Tx_{n_k} \to y$ for some y. Then $\|Tx_n - \lambda x_n\| \to 0 \implies \lambda x_{n_k} \to y$. Since $\|x_{n_k}\| = 1$ for all k, and $\lambda \neq 0$ this shows $\|y\| = \lambda^{-1} \neq 0$. Now, by continuity of T we have also $\lambda Tx_{n_k} \to Ty$. But since $Tx_{n_k} \to y$ by def'n of y, and limits are unique, we have that $\lambda y = Ty$ in other words, y is an eigenvector! ☐

COROLLARY. *(4.15) If T is a compact operator on $\mathcal{H}$and $\lambda \neq 0$ with $\lambda \notin \sigma_p(T)$ and $\overline{\lambda} \notin \sigma_p(T^*)$, then ran$(T - \lambda Id) = \mathcal{H}$ and $(T - \lambda)^{-1}$ is a bounded operator on $\mathcal{H}$.*

PROOF. Since $\lambda \notin \sigma_p(T)$, the preceding proposition implies that there is some constant $c > 0$ such that $\|(T - \lambda)h\| \geq c\|h\|$ for all $h \in \mathcal{H}$. (This very much relies on the fact that T is compact!) This is the essentially estimate that makes everything work.

We claim now that $\text{ran}(T - \lambda Id)$ is closed. Indeed if $(T - \lambda)h_n \to f$ for some f, then we have the estimate $\|h_n - h_m\| \leq c^{-1}\|(T - \lambda)h_n - (T - \lambda)h_m\|$, and so we see that h_n is Cauchy by virtue of the fact that $(T - \lambda)h_n$ is Cauchy. Hence $h_n \to h$ for some h and we conclude that $f = (T - \lambda)h \in \text{ran}(T - \lambda Id)$ after all.

Now, since the range is closed, we can use the identity $\ker A = (\text{ran}A^*)^{\perp}$willy-nilly, and by the funny hypothesis on $\overline{\lambda}$, we have that $\text{ran}(T - \lambda) = [\ker(T - \lambda)^*]^{\perp} = \mathscr{H}$

Finally, to see that $(T - \lambda)^{-1}$ is a bounded operator, one can work a little bit by hand with the operator $\|(T - \lambda)h\| \geq c\|h\|$, or we can just apply the bounded inverse theorem ($T - \lambda$ is a bounded operator and is surjective since its range is all of $\mathcal{H}$ (just proven) and it has trivial kernal (since $\lambda \notin \sigma_p(T)$)) ☐

REMARK. This is sometimes known as **the Fredholm alternative for a compact self-adjonit operator.**

Either: $\lambda \in \sigma_p(A)$ is a true eigenvector of a finite dimensional hilber space OR $T - \lambda I$ is invertable.

REMARK. It will be proven later that if $\lambda \notin \sigma_p(T)$ then $\bar{\lambda} \notin \sigma_p(T^*)$, so this part of the hypothesis is not nessisary.

4.11. The Diagonalization of Compact Self-Adjoint Operators

The main result in this section is:

THEOREM. *(5.1.) [THE SPECTAL THEOREM FOR COMPACT SELF-ADJOINT OPERATORS]*

If T is a compact self-adjoint operator then T has only a countable number of distinct eigenvalues. If $\{\lambda_1, \ldots\}$ are the distrinct nonzero eigenvalues and P_n is the projection of $\mathcal{H}$ onto $\ker(T - \lambda_n)$ then $P_n P_m = P_m P_n = 0$ if $n \neq m$ and each λ_n is real and:

$$T = \sum_{n=1}^{\infty} \lambda_n P_n$$

Where the series converegеs in the sense of the norm topology.

Here are some consequences of this theorem:

COROLLARY. *(5.3) a)* $\ker T = (span_n ranP_n)^{\perp} = (ranT)^{\perp}$
b) Each P_n has finite rank
c) $\|T\| = \sup\{|\lambda_n| : n \geq 1\}$ *and* $\lambda_n \to 0$ *as* $n \to \infty$

PROOF. a) Since $P_n \perp P_m$ for $n \neq m$ we have a Pythagoras type thing $\|Th\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 \|P_n h\|^2$ so $Th = 0 \iff P_n h = 0 \forall n$ and the result follows.

b) Every eigenspace is finite dimensional for compact operators, so this is indeed the case!

c) Can verify that in the right basis, T is indeed diagonal with the λ's on the diagonal so that $\|T\| = \sup\{|\lambda_n| : n \geq 1\}$ is clear (to do this you just have to quotient down to $\overline{\text{ran}(T)}$. Since T is diagonal here, the result follows by 4.6. □

COROLLARY. *(5.4) If T is a compact self-adjoint operator, then there is a sequence μ_n of real numbers and an orthonormal basis e_n for $(\ker T)^{\perp}$ such that for all h,*

$$Th = \sum_{n=1}^{\infty} \mu_n \langle h, e_n \rangle e_n$$

COROLLARY. *If T is a compact self adjoint operator and $\ker T = \{0\}$ then $\mathcal{H}$ is seperable.*

PROPOSITION. *(5.6.) If A is a normal operator and $\lambda \in \mathbb{F}$ then $\ker(A - \lambda) = \ker(A - \lambda)^*$ and $\ker(A - \lambda)$ is a reducing subspace for A (i.e. A issplits up as an operator on $\ker(A - \lambda)$ and another on $\ker(A - \lambda)^{\perp}$)*

PROOF. Since A is normal, so is $A - \lambda$. Hence $\|(A - \lambda)h\| = \|(A - \lambda)^* h\|$ thus $\ker(A - \lambda) = \ker(A - \lambda)^*$. If $h \in \ker(A - \lambda)$ then $Ah = \lambda h \in \ker(A - \lambda)$ by linearity too. On the other hand $A^* h = \bar{\lambda} h \in \ker(A - \lambda)^*$. Therefore, by the equivalent defintion of "reduces" (Namely that $\mathcal{M}$ is invariant for A and A^*, i.e.$A\mathcal{M} \subset \mathcal{M}$ and $A^*\mathcal{M} \subset \mathcal{M}$) we have that $\ker(A - \lambda)$ reduces A. □

REMARK. This proposition is important! The fact that $\ker(A-\lambda)$ reduces A means that we can recursivly work with the spaces $\ker(A-\lambda)$. To see in an example why this is important, take a non-diagonalizable matrix, say $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. The eigenspace $\ker(A-1)$ is e_1 and this DOES NOT reduce A here since there is a non-zero entry in the $e_1 - e_2$ corner of the matrix.

PROPOSITION. *(5.7.) If A is a normal operator and λ, μ are distint eigenvalues, then* $\ker(A-\lambda) \perp \ker(A-\mu)$

PROOF. This is everyones favourite proof from linear algebra! Have that $\langle Ah, g\rangle = \langle h, A^*g\rangle \implies (\lambda-\mu)\langle f, g\rangle = 0$ for every $f, g \in \ker(A-\lambda)$ and $\ker(A-\mu)$ □

PROPOSITION. *(5.8.) If $A = A^*$ and $\lambda \in \sigma_p(A)$ is an eigenvalue, then λ is real*

PROOF. This is everyone's second favourite proof from linear algebra! $\lambda h = Ah = A^*h = \overline{\lambda}h$ □

LEMMA. *(5.9) If $A = A^*$then either $\pm\|T\|$ is an eigenvalue of T.*

PROOF. By Prop 2.13, $\|T\| = \sup\{\langle Th, h\rangle : \|h\| = 1\}$. Hence there is a sequence h_n so that $|\langle Th_n, h_n\rangle| \to \|T\|$, by passing to a subsequence we can assume that $\langle Th_n, h_n\rangle \to \lambda$ where $\lambda = \|T\|$ or $= -\|T\|$. (Here we are almost done....this basically says that λ is an approximate eigenvalue and we know that for compact operators that approximate eigenvalues are all in fact actual eigenvalues.)

Consider now:

$$\begin{aligned} 0 &\leq \|(T-\lambda)h_n\|^2 \\ &= \|Th_n\|^2 - 2\lambda\langle Th_n, h_n\rangle + \lambda^2 \\ &\leq 2\lambda^2 - 2\lambda\langle Th_n, h_n\rangle \to 0 \end{aligned}$$

ANd hence λ is an approximate eigenvalue. Since T is a compact operator, by 4.14 λ is in a true eigenvalue. □

REMARK. In general (for non-Hermitian matrices) this is NOT true. What you can say is that $\|T\| = \sqrt{\lambda_{max}(T^*T)}$, the largest SINGULAR value.

PROOF. (OF THE MAIN SPECTRAL THEOREM) Take $\lambda_1 \in \sigma_p(T)$ so that $\lambda_1 = \|T\|$ and let $\mathcal{E} = \ker(T-\lambda_1)$ and P_1 be the projection on $\mathcal{E}_1$. Let $\mathcal{H}_2 = \mathcal{E}_1^\perp$. By Lemma 5.6, $\mathcal{E}_1$ reduces T, and hence $\mathcal{H}_2$ reduces T as well. Let $T_2 = T|\mathcal{H}_2$ then we verify that T_2 is again a self-adjoint compact operator on $\mathcal{H}_2$.

Now we repeat this procedure on T_2. Find λ_2 so $|\lambda_2| = \|T_2\|$, let $\mathcal{E}_2 = \ker(T_2 - \lambda_2)$. Note that $\{0\} \neq \mathcal{E}_2 \subset \ker(T-\lambda_2)$. Now we can show that $\lambda_1 \neq \lambda_2$ by the fact that $\lambda_1 = \lambda_2$ would contradict that $\mathcal{E}_1 \perp \mathcal{E}_2$.

Put P_2 to to be the projection of $\mathcal{H}$ onto $\mathcal{E}_2$ and $\mathcal{H}_3 = (\mathcal{E}_1 \oplus \mathcal{E}_2)^\perp$. Note that $\|T_2\| \leq \|T\|$ so that $|\lambda_2| \leq |\lambda_1|$.

Repeating this argument inductivly, we get:

i) $|\lambda_1| \geq |\lambda_2| \geq \ldots$

ii) $\mathcal{E}_n = \ker(T-\lambda_n)$ and $|\lambda_{n+1}| = \left\|T|\left(\mathcal{E}_1^\perp \oplus \ldots \oplus \mathcal{E}_n\right)^\perp\right\|$

By i) and the monotone convergence theorem for real numbers, we know there is a limit so that $|\lambda_n| \to \alpha$. This limit is forced to be $\alpha = 0$ because T is a compact operator. Indeed, choose any sequence $e_n \in \mathcal{E}_n$ of norm 1 so that $\|Te_n\| = \lambda_n$. Since T is compact, there is a convergent subsequence Te_{n_k}, since $Te_{n_k} = \lambda_{n_k}e_{n_k} \perp$

$\lambda_{n_j} e_{n_k} = Te_{n_j}$, the only way that this sequence can converge is if $Te_{n_k} \to 0$ (i.e. $\|Te_{n_k} - Te_{n_k}\|^2 = \|Te_{n_k}\|^2 + \left\|Te_{n_j}\right\|^2 = \lambda_{n_k}^2 + \lambda_{n_j}^2 \geq 2\alpha^2$ cannot be a Cauchy sequence unless $\alpha = 0$)

Finally, we claim that $\left\|T - \sum_{j=1}^n \lambda_j P_j\right\| \to 0$. Indeed, $T - \sum_{j=1}^n \lambda_j P_j \equiv 0$ on the space $(\mathcal{E}_1 \oplus \ldots \oplus \mathcal{E}_n)$ by definition, and $T - \sum_{j=1}^n \lambda_j P_j \equiv T$ on the space $(\mathcal{E}_1 \oplus \ldots \oplus \mathcal{E}_n)^\perp$ since the projections here are all orthogonal of each other. Hence $\left\|T - \sum_{j=1}^n \lambda_j P_j\right\| = \left\|T| \left(\mathcal{E}_1 \oplus \ldots \oplus \mathcal{E}_n\right)^\perp\right\| = |\lambda_{n+1}| \to 0$ by the previous argument! $\square$

Banach Spaces

These are notes from Chapter 3 of [**1**].

5.12. Elementary Properties and Examples

Definition. (1.1.) If $\mathcal{X}$ is a vector space over $\mathbb{F}$, a **seminorm** is a function $p : \mathcal{X} \to [0, \infty)$ having the property that:

a) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in \mathcal{X}$

b) $p(\alpha x) = |\alpha| \, p(x)$ for all $\alpha \in \mathbb{F}$ and $x \in \mathcal{X}$

A **norm** has the additional property that:

c) $p(x) = 0 \implies x = 0$

When $\|\cdot\|$ is a norm, $d(x, y) = \|x - y\|$ defines a metric on $\mathcal{X}$.

Definition. (1.2.) A **normed space** is a pair $(\mathcal{X}, \|\cdot\|)$ where $\mathcal{X}$ is a vector space adn $\|\cdot\|$ is a norm. A **Banach space** is a normed space that is complete with respect to the metric defined by the norm.

Proposition. *(1.3.) If $\mathcal{X}$ is a normed space then:*

a) The function $\mathcal{X} \times \mathcal{X} \to \mathcal{X}$ by $(x, y) \to x + y$ is continuous

b) The function $\mathbb{F} \times \mathcal{X} \to \mathcal{X}$ by $(\alpha, x) \to \alpha x$ is continuous

Proof. a) If $x_n \to x$ and $y_n \to y$ then $\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0$.

b) If $\alpha_n \to \alpha$ and $x_n \to x$ then $\|\alpha_n x_n - \alpha x\| \leq \|\alpha_n x_n - \alpha x_n\| + \|\alpha x_n - \alpha x\| \leq |\alpha_n - \alpha| \sup_n \|x_n\| + |\alpha| \, \|x_n - x\| \to 0$. □

Lemma. *If p and q are semi-norms on a vecotr space $\mathcal{X}$, then we write $p \leq q$ to mean that $p(x) \leq q(x)$ for all $x \in \mathcal{X}$. By the scaling property $p(\alpha x) = |\alpha| \, p(x)$, and by continuity, the following are equivalent:*

a) $p(x) \leq q(x)$ for all x

b) $q(x) < 1 \implies p(x) < 1$

c) $q(x) \leq 1 \implies p(x) \leq 1$

d) $q(x) < 1 \implies p(x) \leq 1$

Definition. We say that two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are **equivalent** if they define the same topology on $\mathcal{X}$. (i.e. all limits are the same, all open sets are the same)

Proposition. *(1.5.) If $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if and only if there are positive constants $c, C > 0$ so that:*

$$c \, \|x\|_1 \leq \|x\|_2 \leq C \, \|x\|_1$$

For all $x \in \mathcal{X}$.

PROOF. ($\Rightarrow$) To see that the two topologies are the same, we demonstrate a base of open sets at each point, each open set of which contains an open set from the other topology (the natural base to use for metric spaces is the set of balls centered at a point) From the inequalities in the hypothesis, it is clear that:

$$\begin{aligned} \{x \in \mathcal{X} : \|x - x_0\|_1 \leq \epsilon/C\} &\subset \{x \in \mathcal{X} : \|x - x_0\|_2 \leq \epsilon\} \\ \{x \in \mathcal{X} : \|x - x_0\|_2 \leq c\epsilon\} &\subset \{x \in \mathcal{X} : \|x - x_0\| \leq \epsilon/C\} \end{aligned}$$

($\Leftarrow$) Since $\{x : \|x\|_1 < 1\}$ is an open n'h'd containg 0, it must contain a ball $\{x : \|x\|_2 < r\} \subset \{x : \|x\|_1 < 1\}$. By the proceding lemma, we have that $\|x\|_1 \leq r^{-1} \|x\|_2$. □

EXAMPLE. (1.6.) Let X be any hausdorff space, and let:

$$C_b(X) = \left\{ f : X \to \mathbb{F} : \sup_{x \in X} |f(x)| < \infty \right\}$$

With norm $\|f\| = \sup_{x \in X} |f(x)|$ and pointwise adition and scalare multiplication in the natural way. Then $C_b(X)$ is a Banach space.

PROOF. The only hard thing to check is that $C_b(X)$ is complete.

To do this notice that if f_n is Cauchy in the uniform norm, then $f_n(x)$ is a Cauchy sequence in $\mathbb{F}$ for each $x \in X$. Consequently, since $\mathbb{F}$ is complete, there is a limit, $f(x)$ for each point $x \in X$. We claim now that $f_n \to f$ in the uniform norm. This is an $\epsilon/3$ argument. For any $\epsilon > 0$ take N so large so that $n, m > N \implies \|f_n - f_m\| < \epsilon/3$. Then for any $x \in X$ consider as follows. Find an M_x depending on x so that $|f_n(x) - f(x)| < \epsilon/3$ for all $n \geq M_x$. WOLOG $M_x > N$ and we have that for $n > N$ that $|f_n(x) - f(x)| \leq |f_{M_x}(x) - f(x)| + |f_{M_x}(x) - f_n(x)| < \epsilon/3 + \epsilon/3$. □

PROPOSITION. *(1.7.) If X is a locally compact space, then the space of continuous functions*

$$C_0(X) := \{f \in C_b(X) : \forall \epsilon > 0, \{x \in X : |f(x)| \geq \epsilon\} \text{ is compact}\}$$

Is a closed linear subsapce of $C_b(X)$. Notice that $C_0(\mathbb{R})$ is the set of functions that tend to 0 at $\pm\infty$.

EXAMPLE. (1.8.) The space $L^p(X, \Omega, \mu)$ is a Banach space (this one is a bit trickier to prove....maybe I'll get to it in Bass)

EXAMPLE. (1.10.) Let $n \geq 1$ and let $C^{(n)}[0,1]$ be the collection of functions $f : [0,1] \to \mathbb{F}$ such that f has n continuous derivatives. Define $\|f\| = \sup_{0 \leq k \leq n} \left\{ \sup \left| f^{(k)}(x) \right| : 0 \leq x \leq 1 \right\}$. Then $C^{(n)}[0,1]$ is a Banach space.

EXAMPLE. (1.11.) Let $1 \leq p < \infty$ and let $W_p^n[0,1]$ be the collection of functions $f : [0,1] \to \mathbb{F}$such that f has $n-1$ continuous derivatives, $f^{(n-1)}$ is absolutly continuous, and $f^{(n)} \in L^p[0,1]$. For f in $W_p^n[0,1]$ define:

$$\begin{aligned} \|f\| &= \sum_{k=0}^{n} \left[\int_0^1 \left| f^{(k)}(x) \right|^p \mathrm{d}x \right]^{\frac{1}{p}} \\ &= \sum_{k=0}^{n} \left\| f^{(k)} \right\|_{L^p} \end{aligned}$$

PROPOSITION. *(1.12) If p is a semi-norm on $\mathcal{X}$ then $|p(x)-p(y)| \leq p(x-y)$ and if $\|\cdot\|$ is a norm then $|\|x\|-\|y\|| \leq \|x-y\|$*

PROOF. $p(x)=p(x-y+y) \leq p(x-y)+p(y)$ then rearrange, and do the same trick again with x and y interchanged. □

5.13. Linear Operators on a Normed Space

PROPOSITION. *(2.1.) IF $\mathcal{X}$ and $\mathcal{Y}$ are normed spaces and $A : \mathcal{X} \to \mathcal{Y}$ is a linear map, the follwing are equivalent:*

a) $A \in \mathscr{B}(\mathcal{X},\mathcal{Y})$
b) A is continuous at 0
c) A is continuous at some point
d) There is a positive constant c so that $\|Ax\| \leq c\|x\|$ for all $x \in \mathcal{X}$

PROPOSITION. *$\|A\| = \sup\{\|Ax\| : \|x\| \leq 1\}$ is called the norm of A and $\mathscr{B}(\mathcal{X},\mathcal{Y})$ is a Banach space with this norm as long as $\mathcal{Y}$ is a Banach space.*

PROOF. Suppose A_n is Cauchy. Then for any $x \in \mathcal{X}$, $A_n x$ is a Cauchy sequence in $\mathcal{Y}$ as we have $\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \to 0$. Since $\mathcal{Y}$ is complete, there will be a limit $A_n x \to Ax$. We now claim that A is a linear operator, since $A(x+y) = \lim(A_n(x+y)) = \lim A_n x + \lim A_n y = Ax + Ay$ (scalars are similar). Finally, we verify that A is bounded. By using the fact that A_n is Cauchy, take a subsequence n_k so that $\left\|A_{n_{k+1}} - A_{n_k}\right\| < \frac{1}{2^k}$, and then write that $Ax = \lim_{n\to\infty} A_n x = \lim_{j\to\infty} \sum_{k=1}^{j}\left(A_{n_k} - A_{n_{k-1}}\right)x$ so has $\|Ax\| \leq \sum_{k=1}^{\infty}\left\|\left(A_{n_k} - A_{n_{k-1}}\right)x\right\| \leq \sum_{k=1}^{\infty}\left\|A_{n_k} - A_{n_{k-1}}\right\|\|x\|$, which shows $\|A\| < \sum_{k=1}^{\infty}\left\|A_{n_k} - A_{n_{k-1}}\right\| < \infty$. □

EXAMPLE. (2.2) If (X,Ω,μ) is a $\sigma-$finite measure space and $\phi \in L^\infty(X,\Omega,\mu)$ define $M_\phi : L^p \to L^p$ by $M_\phi f = \phi f$ for all f in L^p. Then $M_\phi \in \mathscr{B}(L^p)$ and $\|M_\phi\| = \|\phi\|_\infty$

EXAMPLE. (2.3.) (Genarlization of the integration kernals example) let (X,Ω,μ) be a positive measure space and suppose $k : X \times X \to \mathbb{F}$ is a $\Omega \times \Omega$ measurable function for which there are constants c_1 and c_2 so that :

$$\begin{aligned} \int |k(x,y)|\,\mathrm{d}\mu(y) &\leq c_1 \text{ for a.e. } x \\ \int |k(x,y)|\,\mathrm{d}\mu(x) &\leq c_2 \text{ for a.e. } y \end{aligned}$$

Then define $K : L^p(\mu) \to L^p(\mu)$ by:

$$(Kf)(x) = \int k(x,y)f(y)\mathrm{d}\mu(y)$$

Then K is a bounded linear operator with $\|K\| \leq c_1^{1/q} c_2^{1/p}$ where $p^{-1}+q^{-1}=1$ (Use Holder instead of Cauchy-Schwarz)

EXAMPLE. (2.4.) If X,Y are compact spaces and $\tau : Y \to X$ is a continuous map, define $A : C(X) \to C(Y)$ by $(Af)(y) = f(\tau(y))$ (i..e $Af = f \circ \tau$) then $A \in \mathscr{B}(C(X),C(Y))$ and $\|A\| = 1$.

Proof. $\|A\| = \sup_{\|f\|_{C(X)}=1} \|Af\|_{C(Y)} = \sup_{\|f\|_{C(X)}=1} \|f \circ \tau\|_{C(Y)} \leq 1$ since $\|f \circ \tau\|_{C(Y)} = \sup_{y \in Y} |f(\tau(y)| \leq \sup_{x \in X} |f(x)| = \|f\|_{C(X)}$. To see the other inequality, construct the right function f which is 1 on some part of the range of τ and ≤ 1 elsewhere. (Such a function should exist by Ursohn-type lemma thinger) $\square$

5.14. Finite Dimensional Normed Spaces

In functional analysis, it is always good to see what significance a concept has for finite dimensional spaces.

Theorem. *(3.1.) If $\mathcal{X}$ is a finite dimensional space over $\mathbb{F}$, then any two norms on $\mathcal{X}$ are equivalent.*

Proof. Fix a (Hamel) basis $\{e_1, \ldots, e_d\}$ for the space and define $\left\|\sum_{i=1}^{d} x_i e_i\right\|_\infty := \max_{1 \leq i \leq d} |x_i|$. We will show that any other norm is equivalent to this norm.

For any x write $x = \sum_j x_j e_j$ and then have by triangle inequality that $\|x\| \leq \sum_j |x_j| \|e_j\| \leq C \|x\|_\infty$ with $C = \sum_j \|e_j\|$. This argument shows moreover that the $\|\cdot\|_\infty$-topology is <u>finer</u> than the $\|\cdot\|$ −topology. (Indeed, if a point x is an interior point for a set S in the $\|\cdot\|$ topology, by scaling by size C, we see that it is an interior point in the $\|\cdot\|_\infty$-topology too. (The releavent logic is $\|x\|_\infty < r/C \implies \|x\| < r$) This means that all the $\|\cdot\|$-open sets are also $\|\cdot\|_\infty$ −open too. i.e. the $\|\cdot\|_\infty$-topology is finer than the $\|\cdot\|$ −topology.)

Now, to see the reverse inclusion/inequality consider as follows. (The crucial idea is that $\|\cdot\|_\infty$ has a COMPACT unit ball) (Here is a Bolazanno-Weirestrass type proof I found on math.stackexchange). Suppose by contradiction, that for all $C > 0, \exists x$ so that $\|x\|_\infty \geq c \|x\|$. Take $c = n$ to get a seqeunce x_n with $\|x_n\|_\infty \geq n \|x_n\|$. By scaling, we can assume WOLOG that $\|x_n\|_\infty = 1$ for all n, i.e. $\|x_n\| \leq \frac{1}{n}$. But then $\|x_n\| \to 0$ shows that $x_n \to 0$ in the $\|\cdot\|$-topology. On the other hand, since the unit ball in $\|\cdot\|_\infty$ is compact, and $\|x_n\| = 1$ for all n, we have by Bolzanno-Weirestrass a convergence subsequence $x_{n_k} \to y$ with $\|y\|_\infty = 1$. Since $\|x_{n_k} - y\| \leq C \|x_{n_k} - y\|_\infty \to 0$ we have that $x_{n_k} \to y$ in $\|\cdot\|$ too. But this is a contradiction as $y \neq 0$!

Conways pf below:

Look at the closed $\|\cdot\|_\infty$ −ball $B = \{x \in \mathcal{X} : \|x\|_\infty \leq 1\}$. Since this is compact in $\|\cdot\|_\infty$ (finite dimensional is used here!), and since the $\|\cdot\|_\infty$-topology is finer than the $\|\cdot\|$-topology, we know that B is compact in the $\|\cdot\|$ −topology too (Think of open subcovers). Moreover, if we restrict our attention from the set $\mathcal{X}$ to the set B, the topologies agree there (I'm not sure why this is true...)

Now the set $A = \{x \in \mathcal{X} : \|x\|_\infty < 1\}$ is a $\|\cdot\|_\infty$ −open set and is $\subset B$, so it is relativly open w.r.t. B in the $\|\cdot\|_\infty$ topology. Hence it is also relativly open in the $\|\cdot\|$ topology. I.e. there is a $\|\cdot\|$-open set U so that $U \cap B = A$. Since $0 \in U$, and U is ,$\|\cdot\|$ −open, we find an $r > 0$ so that $\{x : \|x\| < r\} \subset U$. This is saying:

$$\|x\| < r \text{ and } \|x\|_\infty \leq 1 \implies \|x\|_\infty < 1$$

<u>Claim:</u> $\|x\| < r$ implies that $\|x\|_\infty < 1$

<u>pf:</u> For such an x, write $x = \sum x_j e_j$. Let $\alpha = \|x\|_\infty$ so that $\|x/\alpha\|_\infty = 1$ and $x/\alpha \in B$. Suppose by contradiction now that $\alpha \geq 1$ then $\|x/\alpha\| < r/\alpha \leq r$ and since $\|x/\alpha\|_\infty = 1 \leq 1$ we have then by our choice of r that $\|x/\alpha\|_\infty < 1$ which is a contradiction!

By Lemma 1.4. this shows the other inclusion. $\square$

5.15. Quotients and Products of Normed Spaces

Let $\mathcal{X}$ be a normed space and let $\mathcal{M}$ be a linear manifold in $\mathcal{X}$ and let $Q : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ be the natural quotient map $Qx = x + \mathcal{M}$. We want to make $\mathcal{X}/\mathcal{M}$ into a normed space, so define:

$$\|x + \mathcal{M}\| := \inf \{\|x - y\| : y \in \mathcal{M}\} = \operatorname{dist}(x, \mathcal{M})$$

This is always a semi-norm, and if $\mathcal{M}$ is a closed linear subspace then it is a norm.

THEOREM. *(4.2.) If $\mathcal{M}$ is a closed linear subspace of $\mathcal{X}$ then $\|x + \mathcal{M}\| = dist(x, \mathcal{M})$ is a norm and:*
a) $\|Q(x)\| \leq \|x\|$ for all $x \in \mathcal{X}$ and hence Q is continuous
b) If $\mathcal{X}$ is a Banach space then so is $\mathcal{X}/\mathcal{M}$
c) A subset W of $\mathcal{X}/\mathcal{M}$ is open in $\mathcal{X}/\mathcal{M}$ if and only if $Q^{-1}(W)$ is open in $\mathcal{X}$
d) If U is open in $\mathcal{X}$ then $Q(U)$ is open in $\mathcal{X}/\mathcal{M}$.

PROOF. I'm going to skip the proof for now.
....I'm going to skip some other stuff here too.... □

5.15.1. Products of Normed Spaces.

DEFINITION. Suppose $\mathcal{X}_i$:$i \in I$ is a collection of normed spaces. Define for $1 \leq p < \infty$: =& :

$$\|x\|_{\oplus_p \mathcal{X}_i} := \left[\sum_i \|x(i)\|^p_{\mathcal{X}_i}\right]^{1/p}$$

$$\oplus_p \mathcal{X}_i := \left\{x \in \prod_i \mathcal{X}_i : \|x\|_{\oplus_p \mathcal{X}_i} < \infty\right\}$$

$$\|x\|_{\oplus_\infty \mathcal{X}_i} := \sup_{i \in I} \|x(i)\|_{\mathcal{X}_i}$$

$$\oplus_\infty \mathcal{X}_i := \left\{x \in \prod_i \mathcal{X}_i : \|x\|_{\oplus_\infty \mathcal{X}_i} < \infty\right\}$$

$$\oplus_0 \mathcal{X}_i := \left\{x \in \prod_i \mathcal{X}_i : \|x\|_{\mathcal{X}_n} \to 0 \text{ as } n \to \infty\right\}$$

(The last definition only makes sense when $I = \{1, 2, \ldots\}$. We give $\oplus_0 \mathcal{X}_i$ a norm by treating it as a subspace of $\oplus_\infty \mathcal{X}_i$) The next proposition tells us when this is a Banach space and other things:

PROPOSITION. *(4.4.) Let $\{\mathcal{X}_i : i \in I\}$ be a collection of normed spaces.*
a) $\oplus_p \mathcal{X}_i$ is a normed space and the projections $P_n : \oplus_p \mathcal{X}_i \to \mathcal{X}_n$ is a continuous linear map with $\|P_n(x)\|_{\mathcal{X}_n} \leq \|x\|_{\oplus_p \mathcal{X}_i}$.
b) $\oplus_p \mathcal{X}_i$ is a Banach space if and only if each $\mathcal{X}_n$ is a Banach space
c) Each projection P_n is an open map of $\oplus_p \mathcal{X}_i$ onto $\mathcal{X}_n$.

5.16. Linear Functionals

DEFINITION. A **hyperplane** in $\mathcal{X}$ is a linear manifold $\mathcal{M}$ in $\mathcal{X}$ so that $\dim(\mathcal{X}/\mathcal{M}) = 1$. A little more work gets us to:

PROPOSITION. *(5.1.) a) A linear manifold in $\mathcal{X}$ is a hyperplane if and only if it is the kernal of a non-zero linear functional. b) Two linear functionals have the same kernal if and only if one is a non-zero multiple of the other.*

PROOF. a) If $f : \mathcal{X} \to \mathbb{F}$ is a linear functional and $f \neq 0$ then $\ker f$ is a hyperplane. In fact, f induces an isomorphism between $\mathcal{X}/\ker f$ and $\mathbb{F}$. Conversely, if $\mathcal{M}$ is a hyperplane, let $Q : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ be the natural quotient map and let $T : \mathcal{X}/\mathcal{M} \to \mathbb{F}$be an isomorphism Then $f = T \circ Q$ is a linear functional on $\mathcal{X}$ and $\ker f = \mathcal{M}$.

b) If $\ker f = \ker g$ then take any x_0 with $f(x_0) = 1$. $g(x_0) \neq 0$ since $x_0 \notin \ker f = \ker g$ We claim now that $g(\cdot) = g(x_0)f(\cdot)$. Indeed, $x - f(x)x_0 \in \ker f = \ker g$ for any x, and so $g(x - f(x)x_0) = 0$ for any x. Linearity then gives the result. □

PROPOSITION. *(5.2.) If $\mathcal{X}$ is a normed space and $\mathcal{M}$ is a hyperplane in $\mathcal{X}$, then either $\mathcal{M}$ is closed or $\mathcal{M}$ is dense.*

PROOF. The closure of $\mathcal{M}$ is a linear manifold. Since $\mathcal{M} \subset \text{cl}\mathcal{M}$ and $\dim \mathcal{X}/\mathcal{M} = 1$, either $\text{cl}\mathcal{M} = \mathcal{M}$ or $\text{cl}\mathcal{M} = \mathcal{X}$. □

We will give examples of both these in a second, but first the characterizing theorem is that:

THEOREM. *(5.3.) If $\mathcal{X}$ is a normed space and $f : \mathcal{X} \to \mathbb{F}$ is a linear functional, then f is continuous if and only if* $\ker f$ *is closed. (Otherwise,* $\ker f$ *is dense)*

PROOF. ($\Rightarrow$) If f is continuous, write $\ker f = f^{-1}(\{0\})$ is the pre-image of a closed set and is consequently closed.

($\Leftarrow$) If $\ker f$ is closed, then the quotient map $Q : \mathcal{X} \to \mathcal{X}/\ker f$ is continuous. Let $T : \mathcal{X}/\ker f \to \mathbb{F}$ be an isomortphism and let $g = T \circ Q$. Then g is continuous (its the composition of two continuous functions) and we know from prop 5.1 that $g = \alpha f$ for some constant α. □

EXAMPLE. We now give examples where $\mathcal{M}$ is closed and when it is dense. If we choose $\mathcal{X} = c_0$ (sequence that tend to zero) with the sup norm, then $f(\alpha_1, \alpha_2 \ldots) = \alpha_1$ is a continous linear functional and $\ker f = \{(\alpha_n) : \alpha_1 = 0\}$ is closed.

To make a non-continous linear functional, consider as follows: Take the harmonic sequence $x_0 = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right)$ and then look at the set $\{x_0, e_1, e_2, \ldots\}$ where $e_n(k) = \delta_{nk}$ is the usual basis. This is an independent set. Now extend this to a Hamel basis for all of c_0 (recall a Hamel basis is one where every element is written as a finitie linear combination of basis elements) and define $f(\alpha_0 x_0 + \sum_{n=1}^{\infty} \alpha_n e_n + \sum \beta_i b_i) = \alpha_0$. Notice that $e_n \in \ker f$ for each n and so $\ker f$ is dense.

DEFINITION. For a linear functional f we define:

$$\|f\| = \sup\{|f(x)| : \|x\| \leq 1\}$$

As before, f is continuous if and only if $\|f\| < \infty$. We define:

$$\mathcal{X}^* := \{f : \mathcal{X} \to \mathbb{F} \text{ linear functionals} : \|f\| < \infty\}$$

PROPOSITION. *(5.4.) If $\mathcal{X}$ is a normed linear space then $\mathcal{X}^*$ is a Banach space.*

PROOF. (This essentially goes because $\mathcal{X}^*$ is basically the same as $C_b(\{x : \|x\| < 1\})$ and we know that C_b is a Banach space)

It is easy to check that this the norm $\|\cdot\|$ on functionals is a norm. To see that it is complete, restrict our attention to the unit ball $B = \{x \in \mathcal{X} : \|x\| \leq 1\}$

and define $\rho(f) : B \to \mathbb{F}$ by $\rho(f)(x) = f(x)$ (i.e. $\rho(f)$ is the restriction of the functional f to the closed ball B). Notice that $\rho : \mathcal{X}^* \to C_b(B)$ is a linear isometry. We already know that $C_b(B)$ is complete. Hence to show that $\mathcal{X}^*$ is complete, it suffices to show that $\rho(\mathcal{X}^*)$ is a closed. Indeed, if $f_n \subset \mathcal{X}^*$ and $\rho(f_n) \to g$ for some $g \in C_b(B)$. Define $f : \mathcal{X} \to \mathbb{F}$ by $f(x) = \|x\| g(\|x\|^{-1} x)$ for $\|x\| \neq 0$ and $f(0) = 0$. f is a continuous linear functional because g is bounded. We also see that $\rho(f) = g$, so the fact that $\rho(f_n) \to g$ and $g = \rho(f)$ shows that $\rho(\mathcal{X}^*)$ is closed. □

REMARK. Compare this to the theorem that for $\mathscr{B}(\mathcal{X}, \mathbb{F}) \neq (0)$, we have that $\mathscr{B}(\mathcal{X}, \mathcal{Y})$ is a Banach space if and only if $\mathcal{Y}$ is a Banach space.

THEOREM. *(5.5.) For (X, Ω, μ) a measure space and $1 < p < \infty$ and q s.t. $q^{-1} + p^{-1} = 1$ we define for $g \in L^q$the map $F_g : L^p \to \mathbb{F}$ by:*

$$F_g(f) := \int fg\,d\mu$$

THen $F_g \in (L^p)^*$ and the map $g \to F_g$ is an isometric isomorphism of L^q onto $(L^p)^*$.

5.17. The Hahn-Banach Theorem

DEFINITION. (6.1.) If $\mathcal{X}$ is a vector space, a **sublinear functional** is a function $q : \mathcal{X} \to \mathbb{R}$ such that:

a) $q(x + y) \leq q(x) + q(y)$ for all $x, y \in \mathcal{X}$

b) $q(\alpha x) = \alpha q(x)$ for $x \in \mathcal{X}$ and $\alpha \geq 0$

THEOREM. *(6.2.) Let $\mathcal{X}$ be a vector space over $\mathbb{R}$ and let q be sublinear functional on $\mathcal{X}$. If $\mathcal{M}$ is a linear manifold on $\mathcal{X}$ and $f : \mathcal{M} \to \mathbb{R}$ is a linear functional such that $f(x) \leq q(x)$ for all $x \in \mathcal{M}$ then there exists a linear functional $F : \mathcal{X} \to \mathbb{R}$ such that $F|_{\mathcal{M}} = f$ and $F(x) \leq q(x)$ for all $x \in \mathcal{X}$*

REMARK. The substance of the theorem is not that there exists an extension, but that there exists an extension that is still bounded by q. If one just wanted an extension, then you could just construct one by defining the functional on a Hamel basis.

COROLLARY. *(6.4.) (Complex H-B) If $\mathcal{X}$ is a vector space, let $\mathcal{M}$ be a linear manifold in $\mathcal{X}$ and let $p : \mathcal{X} \to [0, \infty)$ be a seminorm. If $f : \mathcal{M} \to \mathbb{F}$ is a linear functional with $|f(x)| \leq p(x)$ for all $x \in \mathcal{M}$ then there is a linear functional $F : \mathcal{X} \to \mathbb{F}$ sucht ath $F|_{\mathcal{M}} = f$ and $|F(x)| \leq p(x)$ for all $x \in \mathcal{X}$*

....

I think I'm actually not going to type up the rest of this section because all of the information is on the H-B note I made earlier, which is nicely organized. The only theorem I didn's have there that Conway does is this one:

THEOREM. *(6.13) If $\mathcal{X}$ is a normed space and $\mathcal{M}$ is a linear manifold in $\mathcal{X}$ then:*

$$cl\mathcal{M} = \bigcap \{\ker f : \ f \in \mathcal{X}^* \ \textit{and}\ \mathcal{M} \subset \ker f\}$$

PROOF. Let $\mathcal{N}$ be the name of the subspace on the right hand since. We show inclusions both ways.

For any $f \in \mathcal{X}^*$ and $\mathcal{M} \subset \ker f$, since f is continuous we know that $\ker f$ is closed and hence $\mathrm{cl}\mathcal{M} \subset \ker f$. Since this works for any such f then $\mathrm{cl}\mathcal{M} \subset \mathcal{N}$.

We show the other inclusion by contra positive. $x_0 \notin \mathrm{cl}\mathcal{M}$. Then $d = \mathrm{dist}(x_0, \mathcal{M}) > 0$. By the "projection H-B" theorem, we find and f so that $f(x_0) = 1$ and $f|_{\mathcal{M}} = 0$ which shows $x_0 \notin \ker f \supset \mathcal{N}$. □

COROLLARY. *$\mathcal{M}$ is dense in $\mathcal{X}$ if and only if the bounded linear functional on $\mathcal{X}$ that annihilates $\mathcal{M}$ is the zero functional.*

5.18. An Application: Banach Limits

For $x = (x_n) \in c$ the set of sequences with a limit, the operator $L(x) = \lim_{n\to\infty} x_n$ is a linear functional with the folloing properties:
i) $\|L\| = 1$
ii) $L(x) = L(x')$ where $x'_n = x_{n+1}$ is the shifted sequence
iii) $x \geq 0 \implies L(x) \geq 0$
We will show that L extends to a linear operator on all of ℓ^∞ that still has these properties.

THEOREM. *There is a linear operator $L : \ell^\infty \to \mathbb{F}$ so that:*
o) $L(x) = \lim_{n\to\infty} x_n$ for all $x \in c$
i) $\|L\| = 1$
ii) $L(x) = L(x')$ where $x'_n = x_{n+1}$ is the shifted sequence
iii) $x \geq 0 \implies L(x) \geq 0$

PROOF. Let $\mathcal{M} = \{x - x' : x \in \ell^\infty\}$ and notice this a linear manifold. Let $1 = (1, 1, 1, 1 \ldots)$ and check that $\mathrm{dist}(1, \mathcal{M}) = 1$. By the projection H-B theorem, there exists an operator L so that $L(\mathcal{M}) = 0$, $L(1) = 1$ and $\|L\| = \mathrm{dist}(1, \mathcal{M})^{-1} = 1$.

To check that L agrees with limits, it suffices to show that $c_0 \subset \ker L$. To see this, take any sequence $x \in c_0$ and let $x^{(n)} = x'^{\cdots\prime}$ be the sequence shifted n times. Notice that $x^{(n+1)} - x \in \mathcal{M}$ by writing it as a telescoping sum of $x^{(j+1)} - x^{(j)}$. Hence $L(x) = L(x^{(n)}) \leq \left\|x^{(n)}\right\| \to 0$ shows $L(x) = 0$.

To check that L is positive, suppose by contradiction there is a sequence $x \geq 0$ but $L(x) < 0$. WOLOG $\|x\| = 1$ and so $1 \geq x_n \geq 0$ for each n. Have then $\|1 - x\|_\infty \leq 1$ and so $L(1-x) \leq 1$ on the other hand $L(1-x) = 1 - L(x) > 1$ since $L(x) > 0$, contradiction! □

5.19. An Application: Runge's Theorem

Let $\mathbb{C}_\infty$ denote the extended complex plane.

THEOREM. *(Runge's Theorem) Let K be a compact subset of $\mathbb{C}$ and let E be a subset of $\mathbb{C}_\infty \backslash K$ that meets each connected component of $\mathbb{C}_\infty \backslash K$. If f is analytic in a n'h'd of K, then there are rational functions f_n with poles only lying in E so that $f_n \to f$ uniformly on K*

PROOF. I'm skipping the details. The if we let $R(K, E)$ be the closure of the rational functions with poles only in E, then we want to show that $f \in R(K, E)$ for each analytic f. By the geometric fact that $\mathrm{cl}\mathcal{M} = \cap\{\ker \ell\}$ for the right $\ell \in \mathcal{X}^*$ we just have to show that $\ell(f) = 0$ for every $\ell \in \mathcal{X}^*$ for which $\ker \ell \subset R(K, E)$. By Riesz this is the condition that if μ is a measure on K with $\int g \mathrm{d}\mu = 0 \forall g \in R(K, E)$ then $\int f \mathrm{d}\mu = 0$. Having turned the problem into a statement about integrals, the work is more manageable. □

COROLLARY. *(8.5.) If K is compact and $\mathbb{C}\backslash K$ is connected and if f is analytic in a n'h'd of K then there is a sequence of polynomilas that converge to f uniformly on K.*

5.20. An Application: Ordered Vector Spaces

I'm going to skip this section

5.21. The Dual of a Quotient Space and a Subspace

DEFINITION. For a subspace $\mathcal{M} \leq \mathcal{X}$ define $\mathcal{M}^{\perp} \leq \mathcal{X}^*$ by $\mathcal{M}^{\perp} = \{f \in \mathcal{X}^* : \ f(\mathcal{M}) = 0\}$.

THEOREM. *(10.1.) If $\mathcal{M} \leq \mathcal{X}$ then the map $\rho : \mathcal{X}^*/\mathcal{M}^{\perp} \to \mathcal{M}^*$ defined by:*

$$\rho\left(f + \mathcal{M}^{\perp}\right) = f\,|_{\mathcal{M}} \text{ is}$$

an isometric isomorphism.

THEOREM. *If $\mathcal{M} \leq \mathcal{X}$ and $Q : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ is the quotient map, then $\rho(f) = f \circ Q$ defines an isometric isomorphism of $(\mathcal{X}/\mathcal{M})^*$ onto $\mathcal{M}^{\perp}$.*

5.22. Reflexive Spaces

DEFINITION. $\mathcal{X}^{**} = (\mathcal{X}^*)^*$

DEFINITION. For $x \in \mathcal{X}$ we define $\hat{x} \in \mathcal{X}^{**}$ by $\hat{x}(f) = f(x)$. This has $\|\hat{x}\|_{\mathcal{X}^{**}} = \|x\|_{\mathcal{X}}$.

DEFINITION. A space is called **reflexive** if $\mathcal{X}^{**} = \{\hat{x} : x \in \mathcal{X}\}$. i.e. $\mathcal{X}^{**}$ is isometrically isomorhpic to $\mathcal{X}$ by the $\hat{\cdot}$ map.

EXAMPLE. L^p is reflexive, as $(L^p)^{**} = (L^q)^* = L^p$

EXAMPLE. c_0 is not reflexive, as $(c_0)^{**} = \left(\ell^1\right)^* = \ell^{\infty}$.

5.23. The Open Mapping and Closed Graph Theorems

PROPOSITION. *(Added by Mihai) A linear map $A : \mathcal{X} \to \mathcal{Y}$ is an open map $\iff A(B_r(0))$ has non-empty interior*

PROOF. This follows essentially by the translation invariance and scale similarity of the topology on a NVS. For any open set G, write $G = \cup_{x\in G} B_{r_x}(x)$ and then let r_y be the radius of the open ball containing so that $B_{r_y}(0) \subset A\left(B_{r_x}(0)\right)$. Then by translation, we will have: $B_{r_y}(A(x)) \subset A\left(B_{r_x}(x)\right)$ and then can show $A(G) = \cup_{x\in G} B_{r_y}(A(x))$ is an open set. □

THEOREM. *(12.1) (The Open Mapping Theorem) If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $A : \mathcal{X} \to \mathcal{Y}$ is a continuous linear surjection, then $A(G)$ is open in $\mathcal{Y}$ whenever G is open in $\mathcal{X}$.*

I.e. continuous linear surjective maps are open maps.

PROOF. (Sketch) Let $B_r(x)$ denote the unit ball at x. The idea is to show that:
i) $0 \in \text{int cl} A\left(B_r(0)\right)$
This uses that $Y = \cup_n A(B_n(0))$ by "onto" and then Baire to see that these can't all be nowhere dense. This is the harder step.
ii) $\text{cl}\left(A\left(B_{r/2}(0)\right)\right) \subset A\left(B_r(0)\right)$
This just uses the completeness of the space

Combining these, we see that $A(B_r(0))$ has non-empty interior (for 0 is an interior point). The result then follows by the preceding proposition. Lets shore up the details of i) below:

i) Since A is surjective, we know that $Y = \cup_n A\left(B_n(0)\right)$ and now by the Baire category theorem, since Y is a complete space and hence not meager, we know that at least one set $\overline{A\left(B_n(0)\right)}$ has non-empty interior. Say $G \subset \overline{A\left(B_{n_0}(0)\right)}$ is open. Notice that since $B_{n_0}(0)$ is symmetric and convex, since A is linear, we know that $\overline{A\left(B_{n_0}(0)\right)}$ is symmetric and convex too. Hence $\frac{1}{2}G + \frac{1}{2}(-G) \subset \overline{A\left(B_{n_0}(0)\right)}$. But $0 \in \frac{1}{2}G + \frac{1}{2}(-G)$ and this is an open set, so we have then $0 \in \frac{1}{2}G + \frac{1}{2}(-G) \subset$ int cl $A\left(B_{n_0}(0)\right)$ as desired. □

Remark. The Open Mapping Theorem depends very much on the completeness of the space Y; both to use the Baire category theorem AND to see that cl $\left(A\left(B_{r/2}(0)\right)\right) \subset A\left(B_r(0)\right)$

Theorem. *(12.5.) (The Inverse Mapping Theorem) If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $A : \mathcal{X} \to \mathcal{Y}$ is a bounded linear bijection then A^{-1} is bounded.*

Proof. Because A is continuous, linear and surjective, it is an open map by the Open Mapping theorem. This is exactly saying that A^{-1} is a continuous map (using the "preimage of open sets are open" criteria) □

Theorem. *(12.6.) (The Closed Graph Theorem) If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $A : \mathcal{X} \to \mathcal{Y}$ is a linear transformation then:*

$$A \text{ is continuos } \iff \Gamma(A) := \{(x, Ax) \in \mathcal{X} \oplus_1 \mathcal{Y} : x \in \mathcal{X}\} \text{ is a closed set}$$

Remark. $\Gamma(A)$ is called the graph of A, since A is linear it is easily verified that $\Gamma(A)$ is a linear subspace of $\mathcal{X} \oplus_1 \mathcal{Y}$. Recall that the norm $\mathcal{X} \oplus_1 \mathcal{Y}$ is $\|(x, y)\| := \|x\| + \|y\|$.

Lemma. *$\Gamma(T)$ is closed $\iff$ ($x_n \to x$ and $T(x_n) \to y \implies T(x) = y$) $\iff$ ($x_n \to 0$ and $T(x_n) \to y \implies y = 0$)*

Proof. The second $\iff$ is just by linearity, so we prove the first one only:

($\implies$) If $x_n \to x$ and $T(x_n) \to y$ then $(x_n, T(x_n)) \to (x, y)$ since $\Gamma(A)$ is closed we have that $(x, y) \in \Gamma(A)$ and so $y = T(x)$ is forced.

($\impliedby$) Suppose $(x_n, T(x_n)) \to (x, y)$ then we have $x_n \to x$ and $T(x_n) \to y$ and so $T(x) = y$ by hypothesis. Hence $(x, y) = (x, Tx) \in \Gamma(T)$ and we see that $\Gamma(T)$ is closed under limits. □

Proof. (of closed graph theorem)

($\implies$) If A is continuous then the condtion ($x_n \to x$ and $A(x_n) \to y \implies A(x) = y$) follows by continuity of A and see that $\Gamma(A)$ is closed by the lemma.

($\impliedby$) Think of $\Gamma(A)$ as a Banach space (it is a closed subspace of a Banach space) and define $P : \Gamma(A) \to X$ by $(x, Tx) \to x$. It is easily verified that this is bounded and bijective. By the inverse mapping theorem, its inverse is continuous, i.e. $\exists C$ so that $C\|x\| = C\|A(x, Tx)\| \geq C\|(x, Tx)\| = \|x\| + \|Tx\| \implies \|Tx\| \leq (C-1)\|x\|$

(Basically, T is the inverse of the projection map $(x, Tx) \to x$.) □

Remark. The strength of the closed graph theorem is as follows. To show A is continuous you must:

(Without closed graph theorem): If $x_n \to 0$ then must show Ax_n converges AND $Ax_n \to 0$

(With closed graph theorem): If $x_n \to 0$ you may assume that $Ax_n \to y$ converges and have only to show that $y = 0$.

EXAMPLE. If ϕ is a function so that $f\phi \in L^p$ for all $f \in L^p$, show that $\phi \in L^\infty$

PROOF. Let $A : L^p \to L^p$ by $Af = \phi f$. It is possible to verify that $\|A\| = \|\phi\|$ is finite if and only if $\phi \in L^\infty$. To check that A is continuous suppose that $f_n \to 0$ in L^p. We may assume that $\phi f_n \to g$ for some g and we have only to verify that $g = 0$. Since $f_n \to 0$ in L^p we know that $f_n \to 0$ in probability, so there is a subsequence with $f_{n_k} \to 0$ a.s. and we have $g = \lim_k \phi f_{n_k} = 0$ a.s. so indeed $g = 0$.

Notice that without assuming there was such a g, we would be able to show that there was a subsequence with $\phi f_{n_k} \to 0$ but we would be unable to show that $\phi f \to 0$. (We could have used the subsequence-subsubsequence trick for convergence in probability to get around this) □

THEOREM. *(Subsequence-Subsubsequence trick) Let $\mathcal{X}$ be a metric space. Then $x_n \to x$ if and only if for every subsequence x_{n_k} we have a sub-sub-sequence so that $x_{n_{k_m}} \to x$*

PROOF. ($\implies$) is clear. ($\impliedby$) Suppose by contradiciton $x_n \not\to x$. Then there exists $\epsilon_0 > 0$ and a subsequnence n_k so that $\mathrm{d}(x_{n_k}, x) > \epsilon_0$ for all k. But this x_{n_k} has no convergent sub-sub-sequence! □

DEFINITION. (12.8.) If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, an *isomorphism* of $\mathcal{X}$ and $\mathcal{Y}$ is a linear bijection $T : \mathcal{X} \to \mathcal{Y}$ that is a homeomorphism (i.e. it has a continuous inverse). By the inverse mapping theorem all continuous bijections are isomorphisms.

5.24. Complemented Subspaces of a Banach Space

I think I'm going to skip this section.

5.25. The Principle of Uniform Boundedness

THEOREM. *(14.1.) Principle of Uniform Boundednes (or PUB). If $\mathcal{A} \subset \mathscr{B}(\mathcal{X}, \mathcal{Y})$ has the property that:*

$$\forall x \in \mathcal{X}, \sup\{\|Ax\| : A \in \mathcal{A}\} < \infty$$

Then:

$$\sup\{\|A\| : A \in \mathcal{A}\} < \infty$$

(i.e. "If a collection of operators is bounded at each point, then it is uniformly bounded")

PROOF. Let $E_n = \{x \in X : \sup_{\alpha\in\Lambda} \|T_\alpha x\| \le n\}$ so that $E = \cup_n E_n$. Notice also that $E_n = \cap_\alpha \|T_\alpha(\cdot)\|^{-1}[0, n]$ is an intersection of closed sets (because $x \to \|T_\alpha x\|$ is continuous), so E_n is closed. Since E is not 1st category, we know that E cannot be written as a countable union of nowhere dense sets. Hence it must be the case that at least one E_n is not nowhere dense. In other words, $\exists n_0$ so that $E^\circ_{n_0} \ne \emptyset$. Hence $\exists x_0, r$ so that $\overline{B_r(x_0)} \subset E_{n_0}$.

For any x with $\|x\| \leq r$ now, notice that $x_0 + x \in \overline{B_r(x_0)} \subset E_{n_0}$. Hence for such x, we know by definition of E_{n_0} that $\sup_\alpha \|T_\alpha(x_0 + x)\| \leq n_0$. Have then for any $\|x\| \leq r$:

$$\begin{aligned}\sup_\alpha \|T_\alpha x\| &= \sup_\alpha \|T_\alpha(x_0 + x) - T_\alpha(x_0)\| \\ &\leq \sup_\alpha (\|T_\alpha(x_0 + x)\| + \|T_\alpha(x_0)\|) \\ &\leq n_0 + n_0 = 2n_0\end{aligned}$$

So by scaling, we conclude that for any x with $\|x\| \leq 1$ that $\sup_\alpha \|T_\alpha x\| \leq \frac{2n_0}{r}$. Have finally then that $\sup_\alpha \|T_\alpha\| = \sup_\alpha \sup_{\|x\|=1} \|T_\alpha x\| \leq \frac{2n_0}{r} < \infty$. □

Corollary. *(14.3.) If $\mathcal{X}$ is a normed space and $A \subset \mathcal{X}$ then A is a bounded set if and only if for every f in $\mathcal{X}^*$we have that* $\sup\{|f(a)| : a \in A\} < \infty$

Proof. Consider $\mathcal{X}$ as a subset of $\mathcal{X}^{**}$ by the $\hat{}$ map, then it is clear how to apply the PUB. □

Corollary. *(14.4) If $\mathcal{X}$ is a Banach space and $A \subset \mathcal{X}^*$ then A is bounded if and only if for all $x \in \mathcal{X}$ we have* $\sup\{|f(x)| : f \in A\} < \infty$

Proof. This is exactly the uniform boundedness principle with $\mathcal{Y} = \mathbb{F}$. □

Theorem. *(14.6.) The Banach Steinhaus Theorem. If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and A_n is a sequence in $\mathscr{B}(\mathcal{X}, \mathcal{Y})$ with the property that for every $x \in \mathcal{X}$ there exists $y \in \mathcal{Y}$ so that $\|A_n x - y\| \to 0$, then there is a an $A \in \mathscr{B}(\mathcal{X}, \mathcal{Y})$ such that $\|A_n x - Ax\| \to 0$ for every $x \in \mathcal{X}$ and* $\sup \|A_n\| < \infty$

Proof. Let $Ax = \lim_n A_n x$ be the pointwise limit we find. Notice that for each x we have $\|A_n x\| \leq \|A_n x - Ax\| + \|Ax\| \to \|Ax\|$ so we see that for each x , $\|A_n x\|$ is bounded. By the PUB, $\|A_n\|$is uniformly bounded too. Now, to check that A is bounded, write $\|Ax\| \leq \|Ax - A_n x\| + \|A_n x\| \to 0 + \|A_n x\| \leq \sup_n \|A_n\| \|x\|$ so indeed A is continuous too. □

Locally Convex Spaces

These are notes from Chapter 4 of [**1**].

Remark. I'm going to skip around here a bit....some of this stuff is a bit weird

6.26. Elementary Properties and Examples

Definition. (1.1.) A **topological vector space** is a vector space with a topology so that addtion and scalar multiplication are continuous.

Definition. (1.2.) A **locally convex space** is a topological vector space whose topology is defined by a family of seminorms. That is: if $p_1, \ldots, p_n$ are seminorms (i.e. they are like norms by $p(x) = 0$ does not nessisarily mean $x = 0$) then the subbase of the n'hd's at a point $x_0 \in \mathcal{X}$ of the topology are:

$$U_{\epsilon_1,\ldots,\epsilon_n;x_0} = \bigcap_{j=1}^{n} \{x \in \mathcal{X} : p_j (x - x_0) < \epsilon_j\}$$

(i.e. a set U is open iff for all $x_0 \in U$ there exists $\epsilon_1, \ldots, \epsilon_n$ so that $U_{\epsilon_1,\ldots,\epsilon_n;x_0} \subset U$)

Remark. All the seminorms used to define the topology above are automatically continuous since the for every x_0 and$\epsilon > 0$, there is an open n'h'd U of x_0 so that $|p_1(x) - p_1(x_0)| < \epsilon$. (indeed, the n'h'd $U_{\epsilon,1,1,1,\ldots;x_0}$does this)

Remark. I'm skipping a whole bunch of stuff here.

Example. (1.5.) Let $C(X)$ =continuous functions from $X \to \mathbb{F}$. For K compact, define $p_K(f) = \sup\{|f(x)| : x \in K\}$. Then $\{p_K : K \text{ compact in } X\}$ is a family of seminorms that makes $C(X)$ into a locally convex space.

Example. (1.6.) Let G be an open subset of $\mathbb{C}$ and let $H(G)$ be the the set of all analytic functions on G. Define the seminorms of example 1.5. on these functions. It turns out that this is the topology of uniform convergence on compact subsets!

Example. (1.7.) Let $\mathcal{X}$ be a normed space. For each $x^* \in \mathcal{X}^*$define $p_{x^*}(x) = |x^*(x)|$ then p_{x^*} is a semi-norm. If $\mathcal{P} = \{p_{x^*} : x^* \in \mathcal{X}^*\}$, then $\mathcal{P}$ makes $\mathcal{X}$ into a locally convex space. The topology defined on $\mathcal{X}$ by these semi norms is called the **weak topology** and is often denoted by $\sigma(\mathcal{X}, \mathcal{X}^*)$. This is a topology **on X**. In this topology all of the function x^* are continuous.

Example. (1.8.) Let $\mathcal{X}$ be a normed space and for each $x \in \mathcal{X}$, define $p_x : \mathcal{X}^* \to [0, \infty)$ by $p_x(x^*) = |x^*(x)|$. This family of seminorms make s X^* into a topology. This is called the **weak*-topology** It is a topology **on X*.** It is often denoted $\sigma(\mathcal{X}^*, \mathcal{X})$.

DEFINITION. A set A is **convex** if whenever $x, y \in A$ then the line segment $[x, y] = \{tx + (1-t)y : t \in (0,1)\} \subset A$

PROPOSITION. *a) A set A is* ***convex*** *if whenever $x_1, x_2 \ldots x_n \in A$ and $t_1, t_2 \ldots, t_n \in [0,1]$ with $\sum t_i = 1$ has $\sum t_i x_i \in A$*
b) If $\{A_i : i \in I\}$ are all convex, then $\bigcap_i A_i$ is convex.

PROOF. a) by induction, splitting it up into many pairs of points. b) $[x, y] \in A_i$ for each i. □

DEFINITION. (1.10)The **convex hull** of a set A, denoted co(A) is the interesection of all the convex sets that contain A. If $\mathcal{X}$ is a topological vector space, then the **closed convex hull** of A is the intersection of all closed convex subsets of $\mathcal{X}$that contain A; it is denoted by $\overline{\text{co}}(A)$

PROPOSITION. *(1.11) If A is a convex set then:*
a) $\overline{A}$ is a convex set
b) If $a \in A^\circ$and $b \in \bar{A}$ then $(a, b] \subset A^\circ$

PROOF. I'm skiping the proof! □

COROLLARY. *$\overline{co(A)}$ is the closure of $co(A)$*

DEFINITION. A set $A \subset X$ is caleed **balanced** if $\alpha x \in A$ for all $x \in A$ and $|\alpha| \leq 1$. A set A is **absorbing** if for each $x \in \mathcal{X}$ there is an $\epsilon > 0$ so that $tx \in A$ for $0 \leq t < \epsilon$. We say that a set A is **absorbing at** a if for each $x \in \mathcal{X}$ there there is an $\epsilon > 0$ so that $a + tx \in A$ for $0 \leq t < \epsilon$.

REMARK. The condition of being balanced at a is a bit like being an interior point, because the condition is kind of like having a ball of radius ϵ around you, only that the radius can depend on the direction x that was chose. For this reason, sets which are absorbing at every point are not nessarily open.

DEFINITION. For a non-empty convex set V, define the **Mikowski norm** or **Gauge** by:

$$p_V(x) = \inf\{t : t \geq 0, x \in tV\}$$

REMARK. For any set V the set $\{x \in \mathcal{X} : p_V(x) < 1\}$ is a balanced set which is absorbing at each point.

PROPOSITION. *(1.14) If V is a non-empty convex set that is balanced and absorbing at each point, then $V = \{x \in \mathcal{X} : p_V(x) < 1\}$ where p_V is the Minkowski function.*

PROOF. $\{x \in \mathcal{X} : p_V(x) < 1\} \subset V$ is clear. Use the absorbing and balanced properties to prove that $V \subset \{x \in \mathcal{X} : p_V(x) < 1\}$. □

6.27. Metrizable and Normable Locally Convex Spaces

I'm going to skip the details of this section...the main result is that:

PROPOSITION. *(2.1.) A locally convex space is metrizable if and only if it is given by a countable family of seminorms and $\cap_{n=1}^{\infty}\{x : p_n(x) = 0\} = \{0\}$. In this case the metric that works is:*

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{p_n(x-y)}{1 + p_n(x-y)}$$

EXAMPLE. (2.2.) For the locally convex topology we put on $C(X)$ before, it is metrizable iff $X = \bigcup K_n$ whre each K_n is compact and if any compact K is eventually a subset of some K_n.

I'm going to skip the rest.

6.28. Some Geometric Consequence of the Hahn-Banach Theorem

THEOREM. *(3.1.) If $\mathcal{X}$ is a topological vector space and $f : \mathcal{X} \to \mathbb{F}$ is a linear functional then the following are equivalent:*

a) f is continuous

b) f is continuous at 0

c) f is continuous at some point

d) $\ker f$ is closed

e) $x \to |f(x)|$ is a continuous seminorm

And if $\mathcal{X}$ is the locally convex space that is defined by a family of seminors $\mathcal{P}$, then these are equivalent to:

f) $\exists p_1, p_2 \ldots, p_n$ in $\mathcal{P}$ and positive scalars $\alpha_1, \ldots, \alpha_n$ such that:

$$|f(x)| \leq \sum_{k=1}^{n} \alpha_k p_k(x) \ \forall x \in \mathcal{X}$$

PROPOSITION. *(3.2.) Let $\mathcal{X}$ be a topological vector space and suppose that G is an open convex subset of $\mathcal{X}$ that contains the origin. If:*

$$q_G(x) = \inf \{t : t \geq 0 \textit{ and } x \in tG\}$$

then q is a non-negative continuos sublinear functional and $G = \{x : q(x) < 1\}$

REMARK. This is slightly different than proposition 1.14 since that set was balanced and this set is not. The proof is similar to that of Prop 1.14.

THEOREM. *(3.3.) If $\mathcal{X}$ is a topological vector space and G is an open convex non-empty subset of $\mathcal{X}$ that does not contain the origin, then there is a hyperplane $\mathcal{M}$ such that $\mathcal{M} \cap G = \emptyset$*

REMARK. Recall that a hyperplane $\mathcal{M}$ so that $\mathcal{X}/\mathcal{M}$ is one dimensional.

PROOF. The minkowski functional from Proposition 3.2 is the sublinear functional that we will use the H-B theorem to get the result. Pick a point $x_0 \in G$ so that $H := G - x_0$ is an open convex set containing the origin and not containing x_0. By Prop 3.2., then these facts say that $H = \{x : q(x) < 1\}$ and $q(x_0) \geq 1$. On $\text{span}\{x_0\}$ define $f(\alpha x_0) = \alpha q(x_0)$ and so $f \leq q$ on $\text{span}\{x_0\}$ as $f(\alpha x_0) = \alpha q(x_0) = q(\alpha x_0)$ for $\alpha > 0$ and $f(\alpha x_0) = \alpha q(x_0) \leq 0 \leq q(\alpha x_0)$ for $\alpha \leq 0$. Now extend f to all of $\mathcal{X}$ by Hanh Banach.

Now, let $\mathcal{M} = \ker f$ and we will show that this is the hyperplane we want. For $x \in G$ we have $x_0 - x \in H$ and so $f(x_0) - f(x) = f(x_0 - x) \leq q(x_0 - x) < 1 \implies f(x) > f(x_0) - 1 = q(x_0) - 1 \geq 0$. Which shows that f does not vanish anywhere on G.

(The proof for the complex case is simlar....I omit it thought) ☐

DEFINITION. An **affine hyperplane** is a set $\mathcal{M}$ so that for all $x_0 \in \mathcal{M}$, the set $\mathcal{M} - x_0$ is a hyperplane. (it slike a translated hyperplane so that its allowed to not go through 0)

An **affine closed subspace** is a set $\mathcal{M}$ so that for all $x_0 \in \mathcal{M}$, the set $\mathcal{M} - x_0$ is a closed subspace.

REMARK. There is a great advantage inheret in a geometric discussion of real TVS's. Namely, if $f : \mathcal{X} \to \mathbb{R}$ is a nonzero continuos $\mathbb{R}-$linear functional, then the hyperlane $\ker f$ disconnects the space. That is $\mathcal{X}/\ker f$ has two connected components. (This is left as an exercise)

DEFINITION. (3.5.) Let $\mathcal{X}$ be a real topological vector space. A subset S of $\mathcal{X}$ is called an **open half space** if there is a continuous linear functional $f : \mathcal{X} \to \mathbb{R}$ such that $S = \{x \in \mathcal{X} : f(x) > \alpha\}$ for some α. A subset S of $\mathcal{X}$ is called a **closed half space** if there is a continuous linear functional $f : \mathcal{X} \to \mathbb{R}$ such that $S = \{x \in \mathcal{X} : f(x) \geq \alpha\}$ for some α.

DEFINITION. Two subsets A and B of $\mathcal{X}$ are said to be **strictly separated** if they are contained in disjoint open half-spaces; they are **separated** if they are contained in two closed half-spaces whose intersection is a closed affine hyperplane.

PROPOSITION. *(3.6.) Let $\mathcal{X}$ be real topological vector space.*

a) The closure of an open half-space is a closed half-space and the interior of a closed half space is an open half space

b) If $A, B \subset \mathcal{X}$ then A, B are strictly separated if and only if there is a continuous linear functional $f : \mathcal{X} \to \mathbb{R}$ and a real scalar α so that $f(a) > \alpha\ \forall a \in A$ and $f(b) < \alpha\ \forall b \in B$

c) If $A, B \subset \mathcal{X}$ then A, B are separated if and only if there is a continuous linear functional $f : \mathcal{X} \to \mathbb{R}$ and a real scalar α so that $f(a) \geq \alpha\ \forall a \in A$ and $f(b) \leq \alpha\ \forall b \in B$

PROOF. a) is not hard keeping in mind that f is a continuous function.

The ($\Longleftarrow$) direction of b) and c) is clear. Not sure about the other direction....maybe we could use the fact they are disjoint and then look at $G = A - B$ and apply the theorem above to get the functional f....this is essentially the approach of the next theorem. ☐

THEOREM. *(3.7.) If $\mathcal{X}$ is a real topological vector space and A and B are disjoint convex sets with A open, then there is a continuous linear functional $f : \mathcal{X} \to \mathbb{R}$ and real scalar α such that $f(a) < \alpha$ for all $a \in A$ and $f(b) \geq \alpha$ for all b in B. If B is also open then A and B are strictly separated.*

PROOF. Let $G = A - B$ and check that G is convex and does not contain 0. By the previous theorem we have a closed hyperplane so that $\mathcal{M} \cap G = \emptyset$. Take the functional f so that $\mathcal{M} = \ker f$ and for this functional either $f(x) > 0$ for all $x \in G$ or $f(x) < 0$ for all $x \in G$. Suppose that $f(x) > 0$ for all $x \in G$ (the other case is similar) then $f(a-b) > 0$ for all $a \in A$ and $b \in B$ i.e. $f(a) > f(b)$. Taking sups and inf's we get the desired result with $\sup\{f(b) : b \in B\} \leq \alpha \leq \inf\{f(a) : a \in A\}$ ☐

LEMMA. *(3.8.) If $\mathcal{X}$ is a topological vector space, K is a compact subset of $\mathcal{X}$ and V is an open subset of $\mathcal{X}$ sucht that $K \subset V$, then there is an open nhd U of 0 such that $K + U \subset V$.*

THEOREM. *(3.9.) Let $\mathcal{X}$ be a real locally convex space and let A and B be two disjoint closed conves subsets of $\mathcal{X}$. If B is compact, then A and B are strictly seperated.*

THEOREM. *(3.13) If A and B are disjoint closed convex subsets of $\mathcal{X}$, and if B is compact, then there is a continuous functional $f : \mathcal{X} \to \mathbb{C}$ and $\alpha \in \mathbb{R}$ and an $\epsilon > 0$ such that for all $a \in A$ and $b \in B$ we have:*

$$Ref(a) \leq \alpha < \alpha + \epsilon \leq Ref(b)$$

REMARK. I skipped a lot of theorems and some of the optional sections from this chapter.

Weak Topologies

These are notes from Chapter 5 of [**1**].

7.29. Duality

Definition. For a locally convex space $\mathcal{X}$, let $\mathcal{X}^*$ denote the space of all continuous linear functionals on $\mathcal{X}$. This is a vector space. We use the notation $\langle x, x^* \rangle := x^*(x)$ for $x^* \in \mathcal{X}^*$ and $x \in \mathcal{X}$. We will sometimes use $\langle x^*, x \rangle = x^*(x)$ too.

Definition. (1.1.) If $\mathcal{X}$ is a locally convex space, the **weak topology** on $\mathcal{X}$ is the smallest topology so that all the functionals x^* for $x^* \in \mathcal{X}^*$ are continuous. Equivalently, it is defined by the family of seminorms $\{p_{x^*} : x^* \in \mathcal{X}^*\}$ where $p_{x^*}(x) = |\langle x, x^* \rangle|$. This is denoted by "wk" or by $\sigma(\mathcal{X}, \mathcal{X}^*)$

The basic n'h'ds are sets of the form $U_{x_1^*,\ldots,x_n^*;\epsilon,x_0} := \bigcap_{k=1}^n \{x \in \mathcal{X} : |\langle x - x_0, x_k^* \rangle| < \epsilon\}$ and a set U is open if and only if for each $x_0 \in U$ there is a basic open n'h'd $U_{x_1^*,\ldots,x_n^*;\epsilon,x_0} \subset U$. Convergence of nets is characterized by convergence for each linear functional x^*, i..e $x_i \to x_0 \iff \langle x_i, x^* \rangle \to \langle x_0, x^* \rangle$ for all $x^* \in \mathcal{X}^*$.

Definition. The **weak-star topology** on $\mathcal{X}^*$ is the weakest topology so that each functional $\langle \cdot, x \rangle$ for $x \in \mathcal{X}$ is continuous. Equivalently, it is defined by the family of seminorms $\{p_x : x \in \mathcal{X}^*\}$ where $p_x(x^*) = |\langle x, x^* \rangle|$. This is denoted by "wk*" or by $\sigma(\mathcal{X}^*, \mathcal{X})$

The basic n'h'ds are sets of the form $U_{x_1,\ldots,x_n;\epsilon,x_0^*} := \bigcap_{k=1}^n \{x^* \in \mathcal{X} : |\langle x^* - x_0^*, x_k \rangle| < \epsilon\}$ and a set U is open if and only if for each $x_0^* \in U$ there is a basic open n'h'd $U_{x_1,\ldots,x_n;\epsilon,x_0^*} \subset U$. Convergence of nets is characterized by convergence by action on each element x i..e $x_i^* \to x_0^* \iff \langle x_i^*, x \rangle \to \langle x_0^*, x \rangle$ for all $x \in \mathcal{X}$.

Remark. Notice that with these definitions there are two topologies on the set $\mathcal{X}$, the norm topology and the weak topology. We will always use the word "weak" to refer to the weak topology. For example, if we say a set A is closed, that means it is closed in the norm topology. If we wanted to say a set A was closed in the weak topology, we would say "A is weak closed"

Theorem. *(1.4.) If $\mathcal{X}$ is a locally convex space and A is a convex subset of $\mathcal{X}$ then the closure of A is the same as the weak closure of A.*

Proof. Since the weak topology is smaller than the norm topology (i.e. it is a coarser topology), it is clear that cl $A \subset$ wk-cl A.

Conversely, if $x \in \mathcal{X} \backslash \text{cl } A$ then by separation theorems, seperating $\{x\}$ and cl$A$, we know that there is an x^* in $\mathcal{X}^*$ and an α in $\mathbb{R}$ and $\epsilon > 0$ so that:

$$\text{Re}\,\langle a, x^* \rangle \le \alpha < \alpha + \epsilon \le \text{Re}\,\langle x, x^* \rangle$$

for all $a \in \text{cl } A$. Hence $A \subset \text{cl } A \subset B := \{y \in \mathcal{X} : \text{Re}\langle y, x^* \rangle \leq \alpha\}$. But the set B is a weak closed set, since x^* is weak continuous. Thus wk-cl$A \subset B$. Since $x \notin B$, $x \notin$ wk-cl$A$. This shows cl $A^c \subset$ wk-cl A^c , which is the other inclusion. □

COROLLARY. *(1.5.) A convex subset of $\mathcal{X}$ is closed if and only if it is weakly closed.*

REMARK. I skipped the stuff on the polar and prepolar here....

One can reformulate ideas about the Principle of Unifrom boundedness in terms of weak and weak-star topologies.

THEOREM. *(1.10.) If $\mathcal{X}$ is a Banach space and $\mathcal{Y}$ is a normed space and $\mathcal{A} \subset \mathscr{B}(\mathcal{X}, \mathcal{Y})$ is a collection such that for every $x \in \mathcal{X}$ we have $\{Ax : A \in \mathcal{A}\} \subset \mathcal{Y}$ is weakly bounded in $\mathcal{Y}$, then $\mathcal{A}$ is norm bounded in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

7.30. The Dual of a Subspace and a Quotient Space

I'm going to skip the details here, but one of the main results is:

THEOREM. *(2.2.) Let $\mathcal{X}$ be a locally convex space and $\mathcal{M}$ be a closed linear subspace of $\mathcal{X}$. Let $\mathcal{Q}$ be the quotient map from $Q : \mathcal{X} \to \mathcal{X}/\mathcal{M}$.*

Then the map $f \to f \circ \mathcal{Q}$ defines a linear bijection between $(\mathcal{X}/\mathcal{M})^*$ and $\mathcal{M}^\perp = \{x^* \in \mathcal{X}^* : \langle x, x^* \rangle = 0 \forall x \in \mathcal{M}\}$. This bijection is continuous with respect to the weak star topology. If $\mathcal{X}$ is a normed space, then this bijection is an isometry.

7.31. Alaoglu's Theorem

Denote by ball$\mathcal{X} := \{x \in \mathcal{X} : \|x\| \leq 1\}$.

THEOREM. *(3.1.) (Alaoglu's Theorem) If $\mathcal{X}$ is a normed space, then ball$\mathcal{X}^* = \{x^* \in \mathcal{X}^* : \|x^*\| \leq 1\}$ then ball$\mathcal{X}^*$ is weak-star compact.*

PROOF. Let $D = \{\alpha \in \mathbb{F} : |\alpha| \leq 1\}$ be the closed unit disk. Make a copy of this labeled D_x for each x in ball$\mathcal{X}$ and look at the product space $D_\Pi := \prod_{x \in \text{ball}\mathcal{X}} D_x$. This is a compact set by Tychonoff's theorem, since each copy of D is compact. Define $\tau : \text{ball}\mathcal{X}^* \to D_\Pi$ by:

$$\tau(x^*)(x) = \langle x, x^* \rangle$$

I.e. $\tau(x^*)$ is the elmento fo the product wpace D whose x coordinate is $\langle x, x^* \rangle$.

We will show that τ is a homeomorphism from $(\text{ball}\mathcal{X}^*, \text{wk}^*)$ onto the image $\tau(\text{ball}\mathcal{X}^*)$ with the relative topolgogy induced by D_Π and that the range $\tau(\text{ball}\mathcal{X}^*)$ is closed, and hence compact. This will show that $(\text{ball}\mathcal{X}^*, \text{wk}^*)$ is compact, since homeomorphism preserve compactness.

I'm going to skip the actual details here. □

7.32. Reflexivity Revisited

In chapter 3, we said that a Banach space $\mathcal{X}$ was defined to be reflexive if the natural embedding of $\mathcal{X}$ into its double dual $\mathcal{X}^{**}$ was surjective. Recall the map $\hat{} : \mathcal{X} \to \mathcal{X}^{**}$ was given by $x \to \hat{x}$ defined by $\langle x^*, \hat{x} \rangle = \langle x, \hat{x} \rangle$, We showed that this map was an isometry.

If we think of the space $\mathcal{X}^{**}$ as the dual space to $\mathcal{X}^*$, i.e. $\mathcal{X}^{**} = (\mathcal{X}^*)^*$, then we have a weak-star topology on $\mathcal{X}^{**}$; namely its the one characterized by continuity of things from $\mathcal{X}^*$. If we think of $\mathcal{X} \subset \mathcal{X}^{**}$ (by the $\hat{}$ embedding), then the weak-star topology on $\mathcal{X}^{**}$ is exactly the same as the weak topology on $\mathcal{X}$; they are both characterized by the continuity of things from $\mathcal{X}^*$. Being able to think of this in both ways can be useful, for example if $\mathcal{X} = \mathcal{X}^{**}$ is reflexive, in some situations we will be able to apply Alaoglu's theorem to the weak topology of $\mathcal{X}$, since this topology corresponds to the weak-star topology on $\mathcal{X}^{**}$. In the meantime, here are some other results:

PROPOSITION. *(4.1) If $\mathcal{X}$ is a normed space, then ball$\mathcal{X}$ is $\sigma(\mathcal{X}^{**}, \mathcal{X}^*)$ (i.e. the weak-star topology on $\mathcal{X}^{**}$ given by functionals form $\mathcal{X}^*$) dense in ball$\mathcal{X}^{**}$.*

Shorter version: ball$\mathcal{X}$ is dense in ball$\mathcal{X}^{**}$ when $\mathcal{X}^{**}$ is equipped with the weak-start topology.

PROOF. Let B be the closure of ball$\mathcal{X}$ in this topology so that apriori $B \subset$ ball$\mathcal{X}^{**}$ and we desire to show that actuually $B = $ ball$\mathcal{X}^{**}$. Suppose by contradiction that there is an $x_0^{**} \in$ ball$\mathcal{X}^{**} \backslash B$. Then use the Hanh-Banach theorem to seperate the point $\{x_0\}$ from B to find a sperating functional $x^* \in \mathcal{X}^*$ such that:

$$\mathrm{Re}\left(\langle x, x^* \rangle\right) < \alpha < \alpha + \epsilon < \mathrm{Re}\left(\langle x^*, x_0^{**} \rangle\right) \text{ for all } x \in B$$

By scaling the element x^* chosen here, we may assume WOLOG that $\alpha = 1$. Have:

$$\mathrm{Re}\left(\langle x, x^* \rangle\right) < 1 < 1 + \epsilon < \mathrm{Re}\left(\langle x^*, x_0^{**} \rangle\right) \text{ for all } x \in \mathrm{ball}\mathcal{X}$$

$\mathrm{Re}\left(\langle x, x^* \rangle\right) < 1$ shows that $x^* \in$ ball$\mathcal{X}$. But this is a contradiction $1 + \epsilon < \mathrm{Re}\left(\langle x^*, x_0^{**} \rangle\right)$ since $x^* \in$ ball$\mathcal{X}^*$ and $x_0^{**} \in$ ball$\mathcal{X}^{**}$. □

THEOREM. *(4.2.) If $\mathcal{X}$ is a Banach space, the following are equivalent:*

a) $\mathcal{X}$ is reflexive

b) $\mathcal{X}^*$ is reflexive

c) The weak-star topology on $\mathcal{X}^*$ (characterized by action on $\mathcal{X}$) is the same as the weak-topology on $\mathcal{X}^*$ (characterized by functions from $\mathcal{X}^{**}$)

(Conway writes this as $\sigma(\mathcal{X}^*, \mathcal{X}) = \sigma(\mathcal{X}^*, \mathcal{X}^{**})$

d) ball$\mathcal{X}$ is weakly compact.

REMARK. The most important thing to get out of this theorem is that $\mathcal{X}$ is reflexive $\iff$ ball$\mathcal{X}$ is weakly compact.

PROOF. I'm just going to prove that a) $\iff$ d).

a) $\implies$ d) is just Alaoglu's theorem, since the weak topology on $\mathcal{X}$ is the same as the weak-star topology on $\mathcal{X}^{**}$ when $\mathcal{X}$ is reflexive.

d) $\implies$ a) If ball$\mathcal{X}$ is wekaly compact, then it is weakly closed. We have then by the preceding proposition that ball$\mathcal{X} = \overline{\mathrm{ball}\mathcal{X}} = $ ball$\mathcal{X}^{**}$. Hence $\mathcal{X}$ is reflexive! □

DEFINITION. We call a sequence $\{x_n\}$ in $\mathcal{X}$ **weakly Cauchy sequence** if for every $x^* \in \mathcal{X}^*$ we have that $\{\langle x_n, x^* \rangle\}$ is a Cauchy sequence in $\mathbb{F}$. Since $\mathbb{F}$ is complete, and Cauchy sequences are exactly the convergent sequences, you could equally well say that $\langle x_n, x^* \rangle$ is convergent.

THEOREM. *(4.4.) If $\mathcal{X}$ is reflexive, then every weak Cauchy sequence converges weakly. This is called being "weakly sequentially compact".*

PROOF. Say $\{x_n\}$ is the weakly Cauchy sequence in question. Since $\{\langle x_n, x^* \rangle\}$ is Cauchy for each x^*, in particular $\{\langle x_n, x^* \rangle\}$ is bounded for each fixed x^*. By the principle of uniform boundedness, there is a constant M so that $\|x_n\| \leq M$ for all n. (I.e. the operators $\langle x_n, \cdot \rangle$ are bounded at each point x^*, by the PUB there must be a uniform bound. Since $\|\langle x_n, \cdot\rangle\| = \|x_n\|$ this means the x_n are bounded.) Now, since $\mathcal{X}$ is reflexive, the set $\{x \in \mathcal{X} : \|x\| \le M\}$ is weakly compact. Since $\{x_n\}$ is a closed subset of this compact set, we know that there must be a subsequence x_n weakly converging to something, say $x_{n_k} \xrightarrow{weak} x_0$. We know already that $\langle x_n, x^*\rangle$ exists for all x^* (since its Cauchy) so it must be the case that $\langle x_0, x^*\rangle = \lim_{k\to\infty}\langle x_{n_k}, x^*\rangle = \lim_{n\to\infty}\langle x_n, x^*\rangle$for each x^* and we conclude that $x_n \to x_0$ weakly. □

Remark. Not all Banach spaces are weakly sequentially compact. For example, the space $C[0,1]$ with the uniform norm is not. (Remember: its not reflexive, so this is indeed a possibility!) In we take $f_n(t) := \begin{cases} 1 & t = 0 \\ 0 & t \ge \frac{1}{n} \end{cases}$ and linear in the range $[0, 1/n]$ then for any $\mu \in M[0,1] = (C[0,1])^*$ we will have $\int f_n \mathrm{d}\mu \to \mu(\{0\})$ by the Monotone convergence theorem. Since this is convergent, we have by definition that f_n is weakly Cauchy. However, f_n does not converge uniformly to any continuous function in $C[0,1]$!

7.33. Separability and Metrizability

I'm not going to go into too much depth for this section, here are some highlights:

Theorem. *(5.1.) If $\mathcal{X}$ is a Banach space, then ball$\mathcal{X}^*$ is weak-star metrizable if and only if $\mathcal{X}$ is seperable.*

Proposition. *(5.2.) If a sequence in ℓ^1 converges weakly it converges in norm.*

Proof. The main idea is to use that $\left(\ell^1\right)^* = \ell^\infty$ and consequently by Thm 5.1. we have that ballℓ^∞ is weak-star metrizable. One such metrizing is $\mathrm{d}(\phi, \psi) = \sum_j 2^{-j}|\phi(j) - \psi(j)|$. From here you have to do some tricky stuff with the Baire category theorem to get the result. □

7.34. An Application: The Stone-Cech Compactification

Skip!

7.35. The Krein-Milman Theorem

Definition. (7.1.) An **extreme point** of a convex subset A is a point that is never on the interior of a line segment from points in A: it can only be realized as an endpoint. Let ext K be the set of extreme points.

Theorem. *(7.4.) (The Krein-Milman Theorem) If K is a nonempty compact convex subset of a locally convex space $\mathcal{X}$, then ext K is nonempty and $K = \overline{co}\,(\text{ext } K)$*

Fredholm Thoery of Integral Equations

These are notes from Chapter 24 of [**2**].

Suppose $K(x,y)$ is a kernal on $[0,1]^2$, and f is a given function on $[0,1]$. We are interested in finding solutions u to the following system:

$$u(x) + \int_0^1 K(x,y)u(y)\mathrm{d}y = f(x)$$

To begin, lets look at a discretized version of the problem. Fix n and $h = 1/n$, let $K_{ij} = K(ih, jh)$, $f_i = f(ih)$ and $u_j = u(j)$. We now want to solve:

$$u_i + h\sum K_{ij}u_j = f_i$$

The left hand side is the operator, $[\delta_{ij} + hK_{ij}]_{ij}$ which is an $n \times n$ matrix. For this reason, we might be interested in the determinant $\det[\delta_{ij} + hK_{ij}]$. The following claim is an essential tool:

Proposition. *Let A_{ij} be an $N \times N$ matrix. Have:*

$$\begin{aligned} \det[I + hA_{ij}] &= 1 + h\sum_i A_{ii} + \frac{h^2}{2}\sum_{i,j}\det\begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj}\end{bmatrix} + \dots \\ &= 1 + \sum_{k=1}^N h^k \sum_{0\le v_1 < \ldots < v_k \le N-1} \det_{i,j=1}^{k}\left(A_{v_i v_j}\right) \end{aligned}$$

Proof. Let us call $D(h) := \det[I + hA_{ij}]$. $D(h)$ is a polynomial of degree at most N in the variable h, $D(h) = \sum_0^N a_m h^m$, so it suffices to find the coefficients a_m. Using derivatives, we have that:

$$a_m = \frac{1}{m!}\left(\frac{\mathrm{d}}{\mathrm{d}h}\right)^m D(h)\bigg|_{h=0}$$

Label the columns of $I + hA_{ij}$ as $C_j(h)$. Notice that each column $C_j(h)$ has components which are linear in h and also that $C_j(0) = e_j$. Now, think of the determinant as being a linear function of all the columns. Since the derivative is multilinear as a function of the columns C_j, we have the following differentiation rule:

$$\frac{\mathrm{d}}{\mathrm{d}h}\det[C_1(h), C_2(h), \ldots, C_N(h)] = \sum_{k=1}^N \det\left[C_1(h), \ldots, \frac{\mathrm{d}}{\mathrm{d}h}C_k(h), \ldots, C_N(h)\right]$$

(Here a derivative $\frac{\mathrm{d}}{\mathrm{d}h}C_i$ is the vector that we get by taking component-wise derivatives. A skeptical reader could prove this result from the definition of the derivative and using induction, along with the usual "add and subtract" trick that comes up in this type of derivative argument. Hint: Induction hypothesis for $l \le N$

is that: $\lim_{\Delta\to 0}\frac{1}{\Delta}\left(\det\left[C_1(h+\Delta),,\ldots,C_l(h+\Delta),C_{l+1}(h),\ldots,C_N(h)\right]-\det\left[C_1(h),\ldots,C_N(h)\right]\right)=\sum_{k=1}^{l}\det\left[C_1(h),\ldots,\frac{\mathrm{d}}{\mathrm{d}h}C_k(h),\ldots,C_N(h)\right])$ □

Using this rule repeatedly, gives that:

$$\left(\frac{\mathrm{d}}{\mathrm{d}h}\right)^m \det\left[C_1(h),C_2(h),\ldots,C_N(h)\right]=\sum_{k_1,\ldots,k_m=1}^{N}\det\left[C_1(h),\ldots,\frac{\mathrm{d}}{\mathrm{d}h}C_{k_1}(h),\ldots\frac{\mathrm{d}}{\mathrm{d}h}C_{k_m},\ldots C_N(h)\right]$$

In our case, since every component $C_i(h)$ is a linear function of h, taking two derivatives of any column C_k would give a zero-column, and then the determinant from that term would vanish leaving no contribution. For this reason, we only need to consider k_i all distinct. In our effort to evaluate a_m now, we evalute the above at $h=0$. For columns with no derivative we have $C_k(0)=e_k$, and for columns with a derivative we have that $\frac{\mathrm{d}}{\mathrm{d}h}C_k(0)=A_{\cdot k}$ is the column from A. Hence we have:

$$\begin{aligned} m!a_m &= \sum_{k_1\ldots k_m}\det\left[e_1,\ldots,A_{\cdot k_1},\ldots A_{\cdot k_m},\ldots,e_N\right]\\ &= \sum_{k_1\ldots k_m}\det\left[A_{k_ik_j}\right]_{i,j=1}^{m}\end{aligned}$$

If we sort the indeces k_i so that $k_1<\ldots<k_m$ then we pick up a factor of $m!$ which exactly cancels out the $m!$ on the LHS. Plugging back into $D(h)=\sum h^k a_k$ completes the result.

Definition. We define the shorthand:

$$K\begin{pmatrix}x_1 & x_2 & \ldots x_k\\ y_1 & y_2 & \ldots y_k\end{pmatrix}:=\det\left(K(x_i,y_j)\right)_{i,j}\ \ 1\le i,j\le k$$

The formal limit as $n\to\infty$ of the finite sum $D(h):=\det\left[I+hA_{ij}\right]$ as $h=\frac{1}{n}$ is the infinite series:

$$D=\lim_{n\to\infty}D\left(\frac{1}{n}\right)=\sum_{k=0}^{\infty}\frac{1}{k!}\int\int\ldots\int K\begin{pmatrix}x_1 & x_2 & \ldots x_k\\ x_1 & x_2 & \ldots x_k\end{pmatrix}\mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_k$$

This infinite sum is called the **Fredholmf determinant** of the operator $I+K$ i.e. the operator $u(x)\to u(x)+\int_0^1 K(x,y)u(y)\mathrm{d}y$.

Lemma. *This series converges.*

Proof. We use Hadamdards inequality $|\det(C_1,\ldots,C_k)|\le\prod\|C_i\|$. In our case, since $|K(x,y)|\le M$ is bounded, so the length of each column vector in the matrix $K(x_i,x_i)$ is $\le M\sqrt{k}$. Hence by Hadamard, we have $\left|K\begin{pmatrix}x_1 & x_2 & \ldots x_k\\ x_1 & x_2 & \ldots x_k\end{pmatrix}\right|\le M^k k^{k/2}$. Hence the k-th term in the seris is $\le M^k k^{k/2}/k!$ by Stirling's formula this is $\le (Me)^k k^{-k/2}$ which is summable. Hence the thing is absolutly convergent.

One can get a better estimate on the sum if the kernal K is Holder continuous by doing some column operations which don't effect the determinant. □

Let us turn our attention back to the linear system $\left[\delta_{ij}+hK_{ij}\right]_{ij}u_i=f_i$ now. We want to invert this system. The elements of the inverse matrix can be represented by Cramers rule as the determinants of the minor of size 1 less than the order of the system. If you do this, and then pass to the limit $n \to \infty$ as we did before, you get the operator:

$$\begin{aligned} R(x,y) &:= K(x,y) + \int K\begin{pmatrix} x & x_1 \\ y & x_1 \end{pmatrix} \mathrm{d}x_1 + \ldots \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \int \ldots \int K\begin{pmatrix} x & x_1 & \ldots x_k \\ y & x_1 & \ldots x_k \end{pmatrix} \mathrm{d}x_1 \ldots \mathrm{d}x_k \end{aligned}$$

This sum converges uniformly by the same reasoning as before.

Proposition. $R(x,y) + \int K(x,z)R(z,y)dz - DK(x,y) = 0$

Proof. Expand the determinant $K\binom{x\ x_1\ \ldots x_k}{y\ x_1\ \ldots x_k}$ along the first row to get:

$$K\begin{pmatrix} x & x_1 & \ldots x_k \\ y & x_1 & \ldots x_k \end{pmatrix} = K(x,y)K\begin{pmatrix} x_1 & \ldots & x_k \\ x_1 & \ldots & x_k \end{pmatrix} + \sum_{j=1}^{k}(-1)^k K(x,x_j)K\begin{pmatrix} x_1, x_2, \ x_k \\ y, x_1, \ldots \hat{x_j}, \ldots x_k \end{pmatrix}$$

Where $\hat{x_j}$ is the absentee hat indicating that x_j is not there. We now claim that terms appearing in the sum all integrate to the same thing when you do integrate over $\mathrm{d}x_1, \ldots \mathrm{d}x_k$. This is seen by doing row and column swaps to make them all look the same. Have:

$$\begin{aligned} \int \ldots \int (-1)^k K(x,x_j)K\begin{pmatrix} x_1, x_2, \ x_k \\ y, x_1, \ldots \hat{x_j}, \ldots x_k \end{pmatrix} \mathrm{d}x_1 \ldots \mathrm{d}x_k &= \int \ldots \int (-1)^k K(x,z)K\begin{pmatrix} w, x_2, \ldots, z, \ldots x_k \\ y, w, x_2 \ldots x_{j-1}, x_{j+1}, \ldots \end{pmatrix} \\ &= \int \ldots \int (-1)^k(-1)^1 K(x,z)K\begin{pmatrix} z, x_2, \ldots w \ldots \\ y, w, \ldots x_{j-1}, x_{j+1} \end{pmatrix} \\ &= \int \ldots \int (-1)^k(-1)^1(-1)^{k-2} K(x,z)K\begin{pmatrix} z, x_2 \\ y, x_2 \ldots x_j \end{pmatrix} \end{aligned}$$

If we relabel $z = x_1$now and $w = x_k$ then we see these rae all equal! Consequently:

$$\begin{aligned} \int \ldots \int K\begin{pmatrix} x & x_1 & \ldots x_k \\ y & x_1 & \ldots x_k \end{pmatrix} \mathrm{d}x_1 \ldots \mathrm{d}x_k &= \int \ldots \int \left(K(x,y)K\begin{pmatrix} x_1 & \ldots & x_k \\ x_1 & \ldots & x_k \end{pmatrix} + \sum_{j=1}^{k}(-1)^k K(x,x_j)K\begin{pmatrix} x_1, x_2, \ x \\ y, x_1, \ldots \hat{x_j}, \end{pmatrix} \right. \\ &= K(x,y)\left(\int \ldots \int K\begin{pmatrix} x_1 & \ldots & x_k \\ x_1 & \ldots & x_k \end{pmatrix} \mathrm{d}x_1, \ldots \mathrm{d}x_k\right) - k\left(\int \ldots \int K(x,x_1) \right. \end{aligned}$$

If we divide by $k!$ and sum this up now, the LHS becomes $R(x,y)$; the first term on the RHS has a common factor of $K(x,y)$ and the integrals sum to D; by bringin in the $\frac{k}{k!} = \frac{1}{(k-1)!}$ we recognize that we have $R(x_1,y)$ appearing. Hence have:

$$R(x,y) = K(x,y)D - \int K(x,x_1)R(x_1,y)\mathrm{d}x_1$$

As desired. □

Definition. Use the notation $\mathbf{K}$ to be the operator of integration against the kernal K, namely:

$$(\mathbf{K}u)(x) = \int K(x,y)u(y)\mathrm{d}y$$

The equation we are trying to solve is:

$$(\mathbf{I} + \mathbf{K})u = f$$

Notice that:

$$(\mathbf{I}+\mathbf{K})(\mathbf{I}+\mathbf{H}) = \mathbf{I}+\mathbf{L}$$

where $\mathbf{L}$ is the operator with the kernal:

$$L(x,y) = K(x,y) + H(x,y) + \int H(x,z)K(z,y)\mathrm{d}z$$

THEOREM. *If K is a continuous kernal with $D \neq 0$. Then the operator $\mathbf{I}+\mathbf{K}$ is invertable with inverse $\mathbf{I}-D^{-1}\mathbf{R}$.*

PROOF. We have from the proposition that:

$$\begin{aligned} \mathbf{R}+\mathbf{KR}-D\mathbf{K} &= 0 \\ \mathbf{R}+\mathbf{RK}-D\mathbf{K} &= 0 \end{aligned}$$

Since D is assumed to be non-zero this can be rewritten as:

$$\begin{aligned} (\mathbf{I}+\mathbf{K})\left(\mathbf{I}-D^{-1}\mathbf{R}\right) &= \mathbf{I} \\ \left(\mathbf{I}-D^{-1}\mathbf{R}\right)(\mathbf{I}+\mathbf{K}) &= \mathbf{I} \end{aligned}$$

□

The next way to get more information is to instead look at $D(\lambda) = \sum \frac{\lambda^k}{k!}\int\ldots\int K\binom{x_1,\ldots x_k}{x_1,\ldots,x_k}\mathrm{d}x_1\ldots\mathrm{d}x_k$ (so that our D before was $D(1)$) This is what you get if you look at the kernal λK instead of the K. Similarly define $R(x,yl\lambda) = \sum \frac{\lambda^{k+1}}{k!}\int\ldots\int K\binom{x\ x_1,\ldots x_k}{y\ x_1\ldots x_k}\mathrm{d}x_1\ldots\mathrm{d}x_k$. By the estiamtes we had before, you can see that these are entire analytic functions of λ (it is a power series and the estimates we had before show us that $\limsup |C_k|^{1/k} \leq \limsup \left|(Me)^k k^{-k/2}\right|^{1/k} = 0$ so it has an infinite radius of convergence.)

THEOREM. *If K is a continuous kernal such that $D=0$ then the operator $\mathbf{I}+\mathbf{K}$ has a non-trivial null space and is hence not invertible.*

PROOF. For fixed y, let $r(\cdot) = R(\cdot,y)$. Then the fact $R(x,y)+\int K(x,z)R(z,y)\mathrm{d}z + DK(x,y) = R(x,y)+\int K(x,z)R(z,y)\mathrm{d}z = 0$ when we plug in this fixed value of y says:

$$r(x) + \int K(x,z)r(z)\mathrm{d}z = 0$$

i.e. $r(x)$ is in the null space of $\mathbf{I}+\mathbf{K}$. If $r(x)$ is not identically 0, then we are done. The following arguments show that it is not possible that $r(x) \equiv 0$ for every choice of y □

LEMMA. $\int R(x,x;\lambda)dx = \lambda\frac{d}{d\lambda}D(\lambda)$

PROOF. Just write out the power series. □

PROPOSITION. *It is impossible for $R(x,y) \equiv 0$ for all x,y.*

PROOF. Suppose by contradiciton taht $R(x,y)=0$ for all x,y. Then it must have a zero of ∞ order at every point. If $D=0$ then $D(\lambda)$ has a zero at $\lambda=1$. Since $D(\cdot)$ is an analytic function, this is a zero of some finite order m. But then $\int R(x,x;\lambda)\mathrm{d}x = \lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}D(\lambda)$ shows the zeros of R are of finite order. Contradiction!

□

THEOREM. *The complex number κ is an eigenvalue of the integral operator $\mathbf{K}$ if and only if $\lambda = -\frac{1}{\kappa}$ is a zero of $D(\lambda)$*

PROOF. By the above theorems, we had that an operator $\mathbf{I}+\mathbf{K}$ if invertable if $D \neq 0$ and has non-trivial nullspace if $D = 0$. This is saying that 1 is an eigenvaluf of $\mathbf{K}$ if and only if $D = 0$. By applying this to the oeprator $\mathbf{I} + \lambda\mathbf{K}$ we get the result we want, keeping in mind the definiton of $D(\lambda)$. □

THEOREM. *If $\kappa_1, \kappa_2, \ldots$ are the eigenvalues of the integral operator $\mathbf{K}$ whose kernal $K(x, y)$ is Holder continuous in x or y with Holder exponent $> \frac{1}{2}$ then:*

$$\begin{aligned} \int K(x,x)dx &= \sum \kappa_i \\ D &= \prod (1 + \kappa_j) \end{aligned}$$

And the series and the product converge absolutely.

Bounded Operators

These are notes from Chapter 6 of [**3**].

9.36. Topolgies on Bounded operators

Here are three different topologies on $\mathscr{L}(X,Y)$ the space of bounded linear operators from a Banach space X to a Banach space Y.

Name	N'h'd basis at origin	Continuous Functions	Net characterization, $T_n \to T$ iff:
Norm (aka Uniform)	$B_r(0) = \{S : \|S\| < r\}$ (its a Banach space)		$\|T_n - T\| \to 0$ $T_n x \to Tx$ unif for all $\|x\| \leq 1$
Strong Operator Topology	$A_{x_1,\ldots,x_n,\epsilon}(0) := \{S : \|Sx_i\|_Y < \epsilon\}$	$E_x(T) : \mathscr{L}(X,Y) \to Y$ $E_x(T) := T(x)$	$T_n x \to Tx$ for all x
Weak Operator Topology	$A_{\substack{\ell_1,\ldots\ell_n \\ x_1,\ldots,x_n},\epsilon}(0) := \{S : \ell_j(Sx_i) < \epsilon\}$ $\{x_i\} \subset X, \{\ell_j\} \subset Y^*$	$E_{x,\ell}(T) : \mathscr{L}(X,Y) \to \mathbb{C}$ $E_{x,\ell}(T) := \ell(Tx)$	$\ell(T_n x) \to \ell(Tx) \forall x \in X, \ell \in Y^*$

Proposition. *Weak is weaker than strong which is weaker than norm*

Proof. We will show that convergeging in norm implies converging strongly which implies converging weakly. Indeed, the ineqaulies:

$$\|T_n x - Tx\| \leq \|T_n - T\| \, \|x\|$$

Gives the first one, and:

$$|\ell(T_n x) - \ell(Tx)| \leq \|\ell\| \, \|T_n x - Tx\|$$

Gives the second one. □

Example. Here are examples for ℓ^2 that show the different types:
i) $T_n x = \frac{1}{n} x$ has $T_n \to 0$ in the norm topology

ii) $S_n(\xi_1, \xi_2, \ldots) = \left(\underbrace{0, 0, \ldots 0}_{n}, \xi_{n+1}, \xi_{n+2}, \ldots\right)$ has $\|S_n\| = 1$ so $S_n \nrightarrow 0$ in the norm topology, but we do have for any fixed x that $\|S_n x\| \to 0$ meaning that $S_n \to 0$ in the strong operator topology.

iii) $S_n(\xi_1, \xi_2, \ldots) = \left(\underbrace{0, 0, \ldots 0}_{n}, \xi_1, \xi_2, \ldots\right)$ this has $\|S_n x\| = \|x\|$ for every x, so $S_n \nrightarrow 0$ in the strong topology. However for any e_i we have $\langle e_i, S_n x\rangle = 0$ for $n > i$ so we can check that $S_n \to 0$ in the weak operator topology.

Remark. Operators on ℓ^2 can be thought of as infinite $\times$ infitinite matrices with entires $\langle Te_i, e_j\rangle$.

Converging $T_n \to T$ in the norm topology means something like that $\sum_{i,j} |\langle Te_i, e_j\rangle|^2 \to 0$ (Indeed, we roughly have,

$$\begin{aligned} \|T_n x\|^2 &= \sum_{i=1}^{\infty} |\langle T_n x, e_i\rangle|^2 \\ &= \sum_{i,j=1}^{\infty} |x_j|^2 |\langle T_n e_j, e_i\rangle|^2 \end{aligned}$$

So this condition is definetly sufficient. Puttoing x so that $|x_j|^2 = \frac{1}{k}$ $j \leq k$ and 0 otheriwise....I'm spending too much time on this so I'll stop now)

Convering $T_n \xrightarrow{s} 0$ means that $T_n x \to 0$ for each x. It is nessasary and sufficient that $\|T_n e_i\|^2 = \sum_j |\langle T_n e_i, e_j\rangle|^2 \to 0$ for each e_i since any x is approximated by $\sum_{n=1}^{m} x_n e_n$.

Convering $T_n \xrightarrow{w} 0$ means that each $|\langle T_n e_i, e_j\rangle| \to 0$.

Theorem. *(6.1.) Let $\mathscr{L}(\mathscr{H})$ denote the bounded operators on a Hilbert space $\mathscr{H}$. Let T_n be a sequence of bounded operators and suppose that $\langle T_n x, y\rangle$ converges as $n \to \infty$ for each $x, y \in \mathscr{H}$. Then there exists a $T \in \mathscr{L}(\mathscr{H}))$such that $T_n \xrightarrow{w} T$.*

Proof. We first claim that for each x, $\sup_n \|T_n x\| < \infty$. Indeed, for each y we know that $\sup_n |\langle T_n x, y\rangle| < \infty$ so thinking of $\langle T_n x, \cdot\rangle$ as an operator on $\mathscr{H}$, we see that this family is pointwise bounded. By the uniform boundedness principle, it is uniformly bounded. Since the norm of this operator is $\|T_n x\|$ this is exactly saying that $\sup_n \|T_n x\| < \infty$.

Now again by the uniform boundedness principle, it must be that $\sup_n \|T_n\| < \infty$

Define now $B(x,y) = \lim_n \langle T_n x, y\rangle$. This is a sesquilinear form, and it is bounded since $|B(x,y)| \leq \|x\| \|y\| \sup_n \|T_n\|$. By the Riez theorem for Hilbert spaces then, $B(x,y)$ arises as a a bounded linear operator. □

9.37. Adjoitns

Definition. Let X, Y be Banach spaces and T a bounded linear operator from X to Y. The Banach space **adjoint** of $T' : Y^* \to X^*$ is defined by:

$$(T'\ell)(x) := \ell(Tx)$$

Theorem. *The map $T \to T'$ is an isometric isomorphism of $\mathcal{L}(X,Y)$ into $\mathcal{L}(Y^*, X^*)$*

Proof. The fact that it is an isometry comes from using the charaterization $\|x\| = \sup_{\|\ell\| \leq 1} |\ell(x)|$ (This is a consequence of Hanh-Banach). Have:

$$\begin{aligned} \|T\|_{\mathcal{L}(X,Y)} &= \sup_{\|x\|\leq 1} \|Tx\| \\ &= \sup_{\|x\|\leq 1} \sup_{\|\ell\|\leq 1} |\ell(Tx)| \\ &= \sup_{\|\ell\|\leq 1} \left(\sup_{\|x\|\leq 1} |(T'\ell)(x)| \right) \\ &= \|T'\|_{\mathcal{L}(Y^*,X^*)} \end{aligned}$$

We are mostly interested in the case where T is a bounded linear transformation of a Hilber space $\mathcal{H}$ to itself. In a Hilbert space, we know that $\mathcal{H}^* \equiv \mathcal{H}$ by the Riesz theorem, so we may think of $T^* : \mathcal{H} \to \mathcal{H}$. T^*satisfies $\langle x, Ty \rangle = \langle T^* x, y \rangle$. □

THEOREM. *Here are some properties of the adjoint:*
i) $T \to T^*$*is conjugate liear.*
ii) $(TS)^* = S^* T^*$
iii) $(T^*)^* = T$
iv) If T *has a bounded inverse* T^{-1}, *then* T^**has a bounded inverse and* $(T^*)^{-1} = (T^{-1})^*$
e) The map $T \to T^*$*is continuous in the weak and uniform operator topologies, but is only continuous in the strong operator topology if* $\mathcal{H}$ *is finite dimensional.*
f) $\|TT^*\| = \|T\|^2$

PROOF. i)-iv) are routine.

The fact $\langle x, Ty \rangle = \langle T^* x, y \rangle$ shows that if $T_n \xrightarrow{w} T$ then $\langle x, T_n y \rangle \to \langle x, Ty \rangle \iff \langle T_n^* x, y \rangle \to \langle T^* x, y \rangle$ so indeed $T_n^* \xrightarrow{w} T^*$ and we see that the map $T \to T^*$ respects weak limits. The same holds in the uniform topology: if $T_n \to T$ here then $\|T_n - T\| \to 0$ and since $\|T^*\| = \|T\|$, we get that $\|T_n^* - T^*\| \to 0$ too. The shift operator W_n that shifts by n places converges weakly, but not strongly to 0. However, $W_n^* = V_n$ eats the first n componets, and this DOES converge strongly to 0. So $*$ does not repect this convergence.

f) follows since $\|TT^*\| \leq \|T\| \|T^*\| = \|T\|^2$ and conversly we have:

$$\|T^* T\| \geq \sup_{\|x\|=1} \langle x, T^* T x \rangle = \sup_{\|x\|=1} \|Tx\|^2 = \|T\|^2$$

□

DEFINITION. A boudned operator T on a Hilbert space is called **self adjoint** if $T = T^*$

An important class of these are the orthogonal projections:

DEFINITION. If $P \in \mathcal{L}(\mathcal{H})$ and $P^2 = P$, then P is called a **projection.** If in addition, $P = P^*$ then P is called an **orthogonal projection**.

Notice that the range of a projection is always a closed subspace on which P acts like the identity. If in addition P is orthogonal, then P acts like the zero operator on $(\text{ran}P)^\perp$ (Indeed: for $x \in (\text{ran}P)^\perp$ we have $\langle Px, y \rangle = \langle x, Py \rangle = 0$ for any y since $x \in (\text{ran}P)^\perp$is perpendicular to any Py). If $x = y + z$ with $y \in \text{ran}P$ and $z \in (\text{ran}P)^\perp$ is the decomposition guarenteed by the "projection theorem", then $Px = y$. P is called the orthogonal projection onto the subsapce $\text{ran}P$. Thus the projection theorem sets up a one to one correspondence between orthogonal projections and closed subspaces. Since orthogonal projections arise more frequently than non-orthogonal ones, we normally use the word projection to mean orthogonal ones only.

9.38. The Spectrum

Recall that for a linear operator T on $\mathbb{C}^N$, the eigenvalues of T are the complex numbers λ such that the deteriminant of $\lambda I - T$ is equal to zero. The set of such λ is called the spectrum of T. If λ is not an eigenvalue, then $\lambda I - T$ has an inverse since $\det(\lambda I - T) \neq 0$. In infinite dimensions this is more complicated because it is possible that $\lambda I - T$ is invertable but not bounded invertable and other weird things can happen. The spectrum is very important in understanding the operators themselves.

DEFINITION. Let $T \in \mathcal{L}(X)$ a complex number λ is said to be in the **resolvent set** $\rho(T)$ of T if $\lambda I - T$ is a bijection with a bounded inverse. $R_\lambda(T) = (\lambda I - T)^{-1}$ is called the **resolvent** of T at λ. If $\lambda \notin \rho(T)$ then λ is said to be in the **spectrum** $\sigma(T)$ of T.

In other words, $\sigma(T)$ is the set of all λ so that either $\lambda I - T$ is not a bijection or where it does not have a bounded inverse. The inverse mapping theorem, however, guarentees that if $\lambda I - T$ is a continuous bijection then it automatically has a bounded inverse.

DEFINITION. Let $T \in \mathcal{L}(X)$

a) If there is some $x \neq 0$ so that $Tx = \lambda x$ then $x \in \ker(\lambda I - T)$ and so λ must be in the spectrum of T. In this case we call x an **eigenvector** and we call λan **eigenvalue.** The set of all eigenvalues is called the **point spectrum** of T.

b) If $\lambda \in \sigma(T)$ but $\ker(\lambda I - T)$ is trivial, then it must be the case that $\operatorname{ran}(\lambda I - T)$ is not dense in $\mathcal{H}$. (Otherwise we would have that $\lambda I - T$ is invertable). In this case we call the set of such λ the **residual spectrum.**

To study the spcectrum we first develop a theory of functions $\mathbb{C} \to \mathcal{H}$. (The example to keep in mind is the resolvent, $R_\lambda = (\lambda I - T)^{-1}$.) Since we have a norm on $\mathcal{H}$, we can devolp the theory in much the same way as the theory of complex function $f : \mathbb{C} \to \mathbb{C}$.

We say that such a function is **strongly analtic** at some point $z_0 \in \mathbb{C}$ if $\lim_{h \to 0} \frac{x(z_0+h) - x(z_0)}{h}$ exists in $\mathcal{H}$. As in the case of complex valued functions, every strongly analytic function has a norm-convergent taylor series.

A function is called **weakly analytic** if for every linear operator ℓ we have $\ell(x(\cdot))$ is an analytic function in $\mathbb{C}$.

It is a fact that every weakly analytic function is strongly analytic (I am skipping the proof).

THEOREM. *Let X be a Banach space and suppose $T \in \mathcal{L}(X)$. Then the resolvent set $\rho(T)$ is an open subset of $\mathbb{C}$ and $R_\lambda(T)$ is an analytic $\mathcal{L}(X)$-valued function on each component of D. For any $\lambda, \mu \in \rho(T)$, $R_\lambda(T)$ and $R_\mu(T)$ commute with:*

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda) R_\mu(T) R_\lambda(T)$$

PROOF. Ommited for now...you basically just manipulate the power series involved. □

This is called the **first resolvent formula**.

COROLLARY. *For any T, the spectrum of T is not empty.*

PROOF. We can write:

$$R_\lambda(T) = \frac{1}{\lambda}\left(I + \sum_{n=1}^{\infty}\left(\frac{T}{\lambda}\right)^n\right)$$

So as $\lambda \to \infty$ we have $\|R_\lambda\| \to 0$. If $\sigma(T)$ were empty, then $R_\lambda$would be an entire function of λ and then by the Loiville theorem since R_λ is bounded, we would get $R_\lambda = 0$ which is a contradiction. Hence $\sigma(T)$ is not empty. □

The series above is called the **Neumann series** for $R_\lambda(T)$. We also see that $\sigma(T)$ is contained in a disc of radius $\|T\|$ since otherwise the above series is convergent.

THEOREM. $\lim_{n\to\infty} \|T^n\|^{1/n} = r(T) = \sup_{\lambda\in\sigma(T)} |\lambda|$

PROOF. The idea is to prove that the radius of convergence of the Laurent series for R_λ about $\lambda = \infty$ is exactly $r(T)^{-1}$ (Look at the Neumann series). Indeed, the radius of convergence cannot be smaller than $r(T)^{-1}$ since $R_\lambda$is analytic on $\rho(T)$ and $\{\lambda|\lambda > r(T)\} \subset \rho(T)$. On the other hand, the radius of convergence (about ∞) is no more than $r(T)^{-1}$, for it was it would include a point $\lambda \in \sigma(T)$ which is impossible since we know that R_λ is divergent there. $\square$

COROLLARY. *For a Hilbert space* $r(T) = \sup_{\lambda\in\sigma(T)} |\lambda| = \|T\|$.

Facts about the spectrum of an operator

Definition. Let $A \in \mathscr{B}(\mathcal{X})$ be a bounded linear operator on a Banach space $\mathcal{X}$.
The specrum of an operator $\sigma(A)$ is defined to be the set $\sigma(A) = \{\lambda \in \mathbb{C} : \lambda I - A$ does not have a boudned linear
The set $\rho(A) = \mathbb{C} \backslash \sigma(A) = \{\lambda \in \mathbb{C} : I\lambda - A \text{ IS invertable}\}$ is called the resolvent set.
The operator $R_\lambda(A) = (\lambda I - A)^{-1}$ is called the resolvent function.

Remark. Notice that if $\lambda I - A$ is a bijection, then if $\lambda I - A$ is invertable, the inverse is automatically continuous by the inverse mapping theorem. This means that the question of being in the spectrum or not comes down to whether or not $\lambda I - A$ is bijective.

10.39. The resolvent function is analytic and the spectrum is an open, bounded, non-empty set.

Theorem. *(The* ***first resolvent formula****). Let X be a Banach space and suppose $T \in \mathcal{L}(X)$. Then the resolvent set $\rho(T)$ is an open subset of $\mathbb{C}$ and $R_\lambda(T)$ is an analytic $\mathcal{L}(X)$-valued function on each component of D. For any $\lambda, \mu \in \rho(T)$, $R_\lambda(T)$ and $R_\mu(T)$ commute with:*

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda)\, R_\mu(T) R_\lambda(T)$$

Actually the following is more useful:

$$R_\lambda(T) = R_{\lambda_0}(T)\left[I + \sum_{n=1}^{\infty} (\lambda_0 - \lambda)^n R_{\lambda_0}(T)^n\right]$$

Proof. Check by manipulating power series that:

$$\begin{aligned}
R_\lambda(T) &= \frac{1}{\lambda - T} \\
&= \frac{1}{(\lambda - \lambda_0) + (\lambda_0 - T)} \\
&= \frac{1}{\lambda_0 - T}\left(\frac{1}{(\lambda - \lambda_0)(\lambda_0 - T)^{-1} + 1}\right) \\
&= R_{\lambda_0}(T)\left(I + \sum_{k=1}^{\infty} (-1)^k (\lambda - \lambda_0)^k (\lambda_0 - T)^{-k}\right) \\
&= R_{\lambda_0}(T)\left[I + \sum_{n=1}^{\infty} (\lambda_0 - \lambda)^n R_{\lambda_0}(T)^n\right]
\end{aligned}$$

Notice that the radius of convergence of the power series on the RHS is $|\lambda_0 - \lambda| \leq \frac{1}{\left\|R_{\lambda_0}(T)\right\|}$. This shows that $\rho(T)$ is open and that$R_\lambda(T)$ is an analytic function of λ (Since it has a power series expansion around each point.) □

COROLLARY. *The resolvent set is an open set. The spectrum $\sigma(A)$ is hence a closed set.*

PROPOSITION. *The spectrum of a bounded linear operator A is always non-empty.*

PROOF. The resolvent function $R(\lambda) = (\lambda I - A)^{-1}$ is a meromorphic function with singularities only at $\sigma(A)$ (You have to come up with a theory of functions $f : \mathbb{C} \to \mathscr{B}(\mathcal{X})$ but this is exactly analaogous to the theory of holomorphic functions...think of $\mathscr{B}(\mathcal{X})$ as a so called "Banach Algebra") .

Suppose by contradiction $\sigma(A)$ where empty. Then $R(\lambda)$ would be an entire function. But $\|R(\lambda)\| = \left\|(I - \lambda A)^{-1}\right\| = \left\|\frac{1}{\lambda}\left(I + \frac{A}{\lambda} + \frac{A^2}{\lambda^2} + \ldots\right)\right\| \leq \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{\|A\|^k}{\lambda^k} = \frac{1}{\lambda - \|A\|} \to 0$ as $\lambda \to \infty$. Hence $\|R(\lambda)\|$ is a bounded entire function. But then by Louivelle's theorem, it must be a constant. Since $\|R(\lambda)\| \to 0$ it must be that $R(\lambda) \equiv 0$ which is a contradiction! □

DEFINITION. The spectral radius of an operator is:

$$r(A) := \sup_{\lambda \in \sigma(A)} |\lambda|$$

THEOREM. $\lim_{n\to\infty} \|A^n\|^{1/n} = \sup_{\lambda\in\sigma(A)} |\lambda|$

PROOF. The idea is to prove that the radius of convergence of the Laurent series for R_λ about $\lambda = \infty$ is exactly $r(T)^{-1}$ (Look at the Neumann series). Indeed, the radius of convergence cannot be smaller than $r(T)^{-1}$ since $R_\lambda$is analytic on $\rho(T)$ and $\{\lambda|\lambda > r(T)\} \subset \rho(T)$. On the other hand, the radius of convergence (about ∞) is no more than $r(T)^{-1}$, for it was it would include a point $\lambda \in \sigma(T)$ which is impossible since we know that R_λ is divergent there. □

COROLLARY. *$\sigma(A)$ is a bounded set.*

10.40. Subdividing the Spectrum

One can subdivide the spectrum in a few way...I'm going to do it like this:

(1) The **point spectrum (denoted $\sigma_p(T)$)** is the set of true eigenvalues for which there is a non-zero eigenvector x so that $Tx = \lambda x$ (In this case $\ker(\lambda I - T) \supset \{x\}$ and $\lambda I - T$ is not injective.
(2) The **approximate point spectrum (denoted $\sigma_{ap}(T)$)** is the set of points so that $\exists x_n$ of norm 1 so that $Tx_n - \lambda x_n \to 0$. In this case it is possible that $\ker(\lambda I - T) = \emptyset$ but the $\|\lambda I - T\|$ is not bounded from below, i.e. for all c we can find x so that $\|(\lambda I - T)x\| < c\|x\|$ (See the section on the approximate point spectrum).
(3) The **residual spectrum** is the set where $\lambda I - T$ is injective ($\ker(\lambda I - T) = (0)$) but $\lambda I - T$ does not have dense range. (i.e. $\lambda I - T$ is not surjective). If a point is in $\sigma_{ap}(A)\backslash\sigma_p(A)$ i.e. it is in the approximate point spectrum but not an eigenvalue, then it is in the residual spectrum.

10.40.1. The Approximate Point Spectrum.

PROPOSITION. *The following are equivalent:*
a) $\lambda \notin \sigma_{ap}(A)$
b) $\ker(A - \lambda I) = (0)$ *and* $ran(A - \lambda I)$ *is closed*
c) There is a constant $c > 0$ *such that* $\|(A - \lambda I)x\| \geq c\|x\|$ *for all* x

PROOF. a) $\implies$ c): Suppose by contradiction. Then plug in $c = \frac{1}{n}$ to get a sequence x_n such that $\|(A-\lambda I)x_n\| \leq \frac{1}{n}\|x_n\|$. Normalizing x_n to be norm 1 then gives the result.

c) $\implies$ a) Suppose by contradiction $\lambda \in \sigma_{ap}(A)$. Then the sequence x_n of norm 1 $(A-\lambda I)x_n \to 0$ will contradiction the hypothesis c).

c) $\implies$ b): If by contradiction $x \in \ker(A-\lambda I)$ then x would contradict the hypothesis c). If $y_n \in \operatorname{ran}(A-\lambda I)$ and $y_n \to y$ then find x_n so $y_n = (A-\lambda I)x_n$. But then $c\|x_n\| \leq \|(A-\lambda I)x_n\| = \|y_n\|$. Moreover, $c\|x_n - x_m\| \leq \|(A-\lambda I)(x_n - x_m)\| = \|y_n - y_m\|$ so x_n is Cauchy since y_n is. Hence there is a limit $x_n \to x$ and so $y_n = (A-\lambda I)x_n \to (A-\lambda I)x \in \operatorname{ran}(A-\lambda I)$ as desired.

b) $\implies$ c): Let $\mathcal{Y} = \operatorname{ran}(A-\lambda I)$ since this is closed this is a legitimates subspace of $\mathcal{X}$. The map $A-\lambda I : \mathcal{X} \to \mathcal{Y}$ is a bijection since $\ker(A-\lambda I) = (0)$. By the inverse mapping theorem, there is an inverse $B : \mathcal{Y} \to \mathcal{X}$. Have then $\|B(A-\lambda I)x\| = \|x\| \implies \|B\|\|(A-\lambda I)x\| \geq \|x\|$ so c) holds with $c = \|B\|^{-1}$. □

COROLLARY. *Negating each statement gives that the following are equivalent:*

a) $\lambda \in \sigma_{ap}(A)$

b) Either $\ker(A-\lambda) \neq (0)$ (i.e. λ is a true eigenvalue) OR $\ker(A-\lambda)$ is not closed

c) For all $c > 0$*, there exists* x *such that* $\|(A-\lambda I)x\| < c\|x\|$ *(i.e.* A *is "not bounded from below")*

PROPOSITION. $\sigma(A) \subset \sigma_{ap}(A)$

REMARK. Since $\sigma(A)$ is a closed set, $\partial\sigma(A) \subset \sigma(A)$ so the trick is to prove that they are approximate eigenvalues.

PROOF. Let $\lambda \in \sigma(A)$ and let $\rho_n \subset \mathbb{C}\backslash\sigma(A) = \rho(A)$ such that $\rho_n \to \lambda$.

Claim: There is a subsequence n_k so that $\left\|R_{\rho_{n_k}}(A)\right\| = \left\|(A-\rho_{n_k})^{-1}\right\| \to \infty$ as $n \to \infty$

Pf: Suppose by contradiction that $\|R_{\rho_n}(A)\| \leq M$ is bounded. But then by the first resolvent formula we can define $R_\lambda(T)$ by $R_{\rho_{n_0}}(T)\left[I + \sum_{n=1}^{\infty}(\rho_{n_0}-\lambda)^n R_{\rho_{n_0}}(T)^n\right]$ will converge as long as $|\rho_{n_0} - \lambda| < \frac{1}{M}$. This contradicts that λ is not in the resolvent set!

Since the sequence ρ_n was not chosen in any particular way, we now relabel so that ρ_n is the subsequence above with $\left\|(A-\rho_n I)^{-1}\right\| \to \infty$.

Now take x_n of norm 1 so that $\alpha_n = \left\|(A-\rho_n)^{-1}x_n\right\| > \left\|(A-\rho_n)^{-1}\right\| - n^{-1}$. By the claim, $\alpha_n \to \infty$. Put $y_n = \alpha_n^{-1}(A-\rho_n)x_n$ so that $\|y_n\| = 1$ and check that this sequence shows that λ is an approximate eigenvalue.

(Another way to do this is see that if we choose ρ so that $|\rho - \lambda|$ is small and y so that norm $(A-\rho)^{-1}y$ is large compared to norm y, then we will have for $x = (A-\rho)^{-1}y$ that:

$$\begin{aligned}
\|(A-\lambda I)x\| &\leq \|(A-\rho I)x\| + |\lambda-\rho|\,\|x\| \\
&= \|y\| + |\lambda-\rho|\,\|x\| \\
&\leq \frac{1}{M}\left\|(A-\rho)^{-1}y\right\| + |\lambda-\rho|\,\|x\| \\
&= \frac{1}{M}\|x\| + |\lambda-\rho|\,\|x\|
\end{aligned}$$

And we can make M arbitarily large and $|\lambda - \rho|$ arbitarily small by the claim. We then see that $\lambda \in \sigma_{ap}$ since $A - \lambda I$ is not "bounded from below" $\square$

PROPOSITION. *Points $\lambda \in \sigma(A)$ which are poles of A correspond to eigenvalues i.e. $\lambda \in \sigma_p(A)$.*

10.41. The Spectral Theory of Compact Operator

Recall the following facts about compact operators:

PROPOSITION. *In the setting of bounded operators on a hilbert space: A is compact if and only if there is a sequence of finite rank operators A_n so that $\|A_n - A\| \to 0$.*

(The next few results are in the framework of a Hilbert space)

PROPOSITION. *For A a compact operator, and $\lambda \in \sigma_p(T)$ and $\lambda \neq 0$, the eigenspace $\ker(T - \lambda I)$ is fintie dimensional.*

PROOF. Suppose not. Then there is an infinite orthonormal sequence x_n in $\ker(T - \lambda I)$. Since T is compact, there is a subsequence so that Te_{n_k} converges. But this is impossible since $\left\|Te_{n_k} - T_{e_{n_j}}\right\|^2 = \lambda^2 \left\|e_{n_k} - e_{n_j}\right\|^2 = 2\lambda^2 > 0.$ $\square$

PROPOSITION. *Suppose T is compact. If $\lambda \neq 0$ and $\lambda \in \sigma_{ap}(T)$ is an approximate eigenvalue, then $\lambda \in \sigma_p(T)$ is a true eigenvalue. eigenvalues i.e $\sigma_{ap}(T) = \sigma_p(T)$.*

PROOF. We will show that if h_n are unit vectors so that $\|(T - \lambda I)h_n\| \to 0$ then there exists h so that $\|(T - \lambda I)\, h\| = 0$. Since T is compact, Th_n has a convergent subsequence, say $Th_n \to g$. We claim now that $\lambda h_n \to g$ indeed, $\|\lambda h_n - g\| = \|(T - \lambda I)h_n - (Th_n - g)\| \leq \|(T - \lambda I)h_n\| + \|Th_n - g\| \to 0 + 0$. Since $\lambda \neq 0$ we have $(T - \lambda I)\left(\frac{1}{\lambda} g\right) = \lim_{n\to\infty} (T - \lambda I)\, h_n = 0$ so $\frac{1}{\lambda} g$ is an eigenvector for T. $\square$

COROLLARY. *If T is a compact operator on $\mathcal{H}$ and $\lambda \neq 0$ and $\lambda \notin \sigma_p(T)$ then ran$(T - \lambda) = \mathcal{X}$ and $(T - \lambda I)^{-1}$ is a bounded operator on $\mathcal{X}$.*

PROOF. By the preceding proposition, λ is not an approximate eigenvalue, i.e. $\lambda \notin \sigma_{ap}(T)$. Hence $T - \lambda I$ is bounded from below, i.e. there is a constant c such that $\|(T - \lambda I)\, x\| \geq c\, \|x\|$ and $\operatorname{ran}(T - \lambda I)$ is closed...... we do some more work to show that actually $\operatorname{ran}(T - \lambda I)$ is all of $\mathcal{X}$ and the inverse is then bounded by $\left\|(T - \lambda I)^{-1}\right\| \leq c^{-1}$. $\square$

These lead to the following Spectral Theorem for Compact operators:

THEOREM. *(Riesz) If $\dim \mathcal{X} = \infty$ and $A \in \mathscr{B}_0(\mathcal{X})$ is a compact operator, then only of the following possibilities occur:*
a) $\sigma(A) = \{0\}$
b) $\sigma(A) = \{0, \lambda_1, \ldots, \lambda_n\}$ where each eigenspace is finite dimensional.
c) $\sigma(A) = \{0, \lambda_1, \ldots\}$ where each eigenspace is finite dimensional and 0 is the ONLY limit point of the set $\{\lambda_k\}$ i.e $\lambda_k \to 0$ as $k \to \infty$.

Here is another way to state it:\

THEOREM. *For a compact operator A:*

i) Every nonzero $\lambda \in \sigma(A)$ is an eigenvalue of A.

ii) For all nonzero $\lambda \in \sigma(A)$, there exist m such that $\ker(\lambda I - A)^m = \ker(\lambda I - A)^{m+1}$ *and this subspace is finite dimensional.*

iii) The eigenvalues can only accumulate at 0. If the dimension of ran(A) is not finite, then $\sigma(A)$ must contain 0.

iv) $\sigma(A)$ is countable.

v) Every nonzero $\lambda \in \sigma(A)$ is a pole of the resolvent function $R_\lambda(A)$

Bibliography

[1] J.B. Comway. *A Course in Functional Analysis*. Graduate Texts in Mathematics. Springer New York, 1994.

[2] P.D. Lax. *Functional analysis*. Pure and applied mathematics. Wiley, 2002.

[3] M. Reed and B. Simon. *Methods of modern mathematical physics: Functional analysis*. Methods of Modern Mathematical Physics. Academic Press, 1972.