FUNCTIONAL ANALYSIS READING COURSE SUMMER 2020

JAMES MARSHALL REBER JAMESMARSHALLREBER@GMAIL.COM
THOMAS O'HARE MMMSTOHARE@GMAIL.COM

Contents

Sections correspond to different texts, subsections correspond to chapters. Problems may have different formatting when compared to the book but they should still contain the same content. Each problem should be signed with who did it and with references.

As far as reading goes, we followed [3]. We covered chapters 1-4, 10, and part of 11. There was a week where we covered fixed points theorems (Thomas covered Kakutani and Markov-Kakutani, James covered Ryll-Nardzewski).

This document was compiled by James. Any typos or major mistakes are James' fault.

1. Rudin Solutions

1.1. Chapter 1.

Problem 1 (Rudin 1.1, James)**.** Suppose X is a vector space. All sets mentioned below are understood to be subsets of X. Prove the following statements from the following axioms:

(i) To every pair of vectors x and y corresponds a vector $x+y$ in such a way that

$$x+y=y+x, \text{ and } x+(y+z)=(x+y)+z.$$

(ii) X contains a unique vector 0 such that $x+0=x$ for all $x\in X$.

(iii) For all $x\in X$, there exists a unique vector $-x$ so that $x+(-x)=0$.

(iv) For every $\alpha\in\Phi$ and $x\in X$ there is a vector αx such that

$$1x=x, \text{ and } \alpha(\beta x)=(\alpha\beta x)$$

and such that the two distributive laws

$$\alpha(x+y)=\alpha x+\alpha y \text{ and } (\alpha+\beta)x=\alpha x+\beta x.$$

(1) If $x\in X$ and $y\in X$ there is a unique $z\in X$ so that $x+z=y$.
(2) $0x=0=\alpha 0$ if $x\in X$ and α a scalar.
(3) $2A\subset A+A$.
(4) A is convex if and only if $(s+t)A=sA+tA$ for all positive scalars s and t.

(5) Every union (and intersection) of balanced sets is balanced.
(6) Every intersection of convex sets is convex.
(7) If Γ is a collection of convex sets that is totally ordered by set inclusion, then the union of all members of Γ is convex.
(8) If A and B are convex, then so is $A+B$.
(9) If A and B are balanced, so is $A+B$.
(10) Show that (6), (7), and (8) hold with subspaces in place of convex sets.

Proof. (1) We first establish existence. We claim that $z = y - x$ is such a vector. This follows, since

$$x + z = x + (y - x) = (x + y) - x = (y + x) - x = y + (x - x) = y + 0 = y.$$

We claim that this z is unique. If there is another such z', we have that

$$x+z = x+z' \implies (-x)+(x+z) = (-x)+(x+z') \implies ((-x)+x)+z = ((-x)+x)+z' \implies z = z'.$$

(2) First we note that $0x = 0$. This follows, since

$$x = (0 + 1)x = 0x + 1x = 0x + x,$$

so

$$0 = x + (-x) = (0x + x) + (-x) = 0x + (x + (-x)) = 0x + 0 = 0x.$$

Next, we claim that $\alpha 0 = 0$. This follows, since for some $x \in X$, we have

$$\alpha 0 = \alpha(x + (-x)) = \alpha x + \alpha(-x) = \alpha x + (-\alpha)x = (\alpha + (-\alpha))x = 0x = 0.$$

(3) We claim that $2A \subset A + A$. Notice that for $A \subset X$, we have that

$$2A = \{2x : x \in A\} = \{x + x : x \in A\} \subset \{x + y : x, y \in A\} = A + A.$$

Here we implicitly use $2 = 1 + 1$, so $2x = (1 + 1)x = x + x$.

(4) ($\implies$): Assume that A is convex. We have that for all $0 \le t \le 1$ that

$$tA + (1 - t)A \subset A.$$

Notice that $(s + t)x = sx + tx$, so we see that for all subsets A

$$(s + t)A = \{(s + t)x : x \in A\} = \{sx + tx : x \in A\} \subset \{sx + ty : x, y \in A\} = sA + tA.$$

It suffices then to show that $sA + tA \subset (s + t)A$. If s and t are 0, then the result clearly follows, since $0A = 0$. Assume at least one of s or t is non-zero. Then we can multiply both sides of the above by $1/(s + t)$. After relabeling, this implies that $0 \le s, t \le 1$ so that $s + t = 1$. Hence, we have $s = 1 - t$, so we can rewrite this as

$$tA + (1 - t)A = tA + sA \subset (s + t)A = A$$

by convexity. Thus, the inequality holds.
($\impliedby$): If, for all scalars s and t, we have $sA + tA = (s + t)A$, then in particular for scalars s and t with $0 \le s, t \le 1$ and with $s + t = 1$ we have $sA + tA = A$, or in other words, for $0 \le t \le 1$ we have

$$tA + (1 - t)A \subset A.$$

So A is convex.

(5) We first show that every union of balanced sets is balanced. Recall that a set $U \subset X$ is balanced if for all scalars α with $|\alpha| \le 1$, we have $\alpha U \subset U$. We first show that if $\{U_n\}$ is a family of subsets of X, then for scalars α we have

$$\alpha \bigcup U_n = \bigcup \alpha U_n.$$

Notice that

$$\alpha\bigcup U_n = \{\alpha x : x \in U_n \text{ for some } n\} = \{\alpha x : \alpha x \in \alpha U_n \text{ for some } n\} = \bigcup \alpha U_n.$$

Hence, if $|\alpha| \le 1$ and $\{U_n\}$ is a family of balanced subsets of X, then

$$\alpha\bigcup U_n = \bigcup \alpha U_n,$$

and $\alpha U_n \subset U_n$ for each n, so

$$\alpha\bigcup U_n \subset \bigcup U_n.$$

Similarly, if $\{U_n\}$ is a family of subsets of X and α a scalar, we have

$$\alpha\bigcap U_n = \{\alpha x : x \in U_n \text{ for all } n\} = \{\alpha x : \alpha x \in \alpha U_n \text{ for all } n\} = \bigcap \alpha U_n,$$

so if the U_n are balanced and $|\alpha| \le 1$, we have

$$\alpha\bigcap U_n = \bigcap \alpha U_n \subset \bigcap U_n.$$

Thus, $\bigcup U_n$ and $\bigcap U_n$ are balanced.

(6) Let $\{U_n\}$ be a family of convex sets. Then we see that

$$t\bigcap U_n + (1-t)\bigcap U_n \subset \bigcap (tU_n + (1-t)U_n) = \bigcap U_n.$$

So the family is convex. Notice that a union of convex sets is not necessarily convex; the sets $\{0\}$ and $\{1\}$ in $\mathbb{R}$ are convex, but $\{0\}\bigcup\{1\}$ is not convex.

(7) We need to show that

$$t\bigcup_{V\in\Gamma} V + (1-t)\bigcup_{V\in\Gamma} V \subset \bigcup_{V\in\Gamma} V.$$

Let x be in the left hand side. Then $x = ty + (1-t)z$ for $y, z \in \bigcup_{V\in\Gamma} V$. Since there is a total ordering on Γ, we can find U sufficiently large so that $y, z \in U$. U is convex by construction, so $ty + (1-t)z \in U$, and hence $x \in U$. So $x \in \bigcup_{V\in\Gamma} V$. Since the choice of x was arbitrary, we have that the "inequality" holds, so the union is convex.

(8) Suppose A and B are convex. Note that

$$A + B = \{x + y : x \in A, y \in B\}.$$

Fix $0 \le t \le 1$. The goal is to show

$$t(A+B) + (1-t)(A+B) \subset A + B.$$

This follows by the distributive property, though. For α a scalar, we have

$$\alpha(A+B) = \{\alpha(x+y) : x \in A, y \in B\} = \{\alpha x + \alpha y : x \in A, y \in B\} = \alpha A + \alpha B,$$

so using commutativity, we have

$$t(A+B)+(1-t)(A+B) = tA+tB+(1-t)A+(1-t)B = (tA+(1-t)A)+(tB+(1-t)B) \subset A+B.$$

(9) By the same argument as in (8), we see that for $|\alpha| \le 1$,

$$\alpha(A+B) = \alpha A + \alpha B \subset A + B,$$

so $A + B$ is balanced.

(10) Let $\{V_n\} \subset X$ be vector subspaces. This means that V_n is closed under addition and scalar multiplication for all n. We wish to show that $V_n \cap V_m$ is a vector subspace for all m, n. Note that being closed under addition means

$$V_n + V_n \subset V_n,$$

and being closed under scalar multiplication implies for all scalars α, we have

$$\alpha V_n \subset V_n.$$

Thus, we see

$$\bigcap V_m + \bigcap V_m \subset V_n \text{ for all n},$$

so

$$\bigcap V_m + \bigcap V_m \subset \bigcap V_m.$$

An analogous argument applies for scalar multiplication. Hence, $\bigcap V_n$ is a vector subspace.

The same kind of argument as in (7) tells us that the union of totally ordered vector subspaces is a vector subspace.

We see that

$$(V_1 + V_2) + (V_1 + V_2) = (V_1 + V_1) + (V_2 + V_2) \subset V_1 + V_2,$$
$$\alpha(V_1 + V_2) = \alpha V_1 + \alpha V_2 \subset V_1 + V_2,$$

so $V_1 + V_2$ is a vector subspace if V_1 and V_2 are vector subspaces.

□

Problem 2 (Rudin 1.2, James)**.** The *convex hull* of a set A in a vector space X is the set of all convex combinations of members of A; that is, the set of all sums

$$t_1 x_1 + \cdots + t_n x_n,$$

in which $x_i \in A$, $t_i \geq 0$, $\sum t_i = 1$, n is arbitrary. Denote the convex hull by $\mathrm{Conv}(A)$. Prove that the convex hull of A is convex, and it is the intersection of all convex sets that contain A.

Proof. Fix $0 \leq t \leq 1$. We need to show that

$$t\mathrm{Conv}(A) + (1-t)\mathrm{Conv}(A) \subset \mathrm{Conv}(A).$$

Taking an element in the left hand side, we have that it is in the form

$$t\left[\sum_{j=1}^{m} t_j a_j\right] + (1-t)\left[\sum_{k=1}^{n} t_k a_k\right],$$

where $\sum_{j=1}^m t_j = 1$, $\sum_{k=1}^n t_k = 1$. We see that

$$t\sum_{j=1}^{m} t_j + (1-t)\sum_{k=1}^{n} t_k = t + (1-t) = 1,$$

so after relabeling we have that the element is in the form

$$\sum_{l=1}^{m+n} t_l a_l,$$

where $a_j \in A$ for $1 \leq j \leq m+n$ and $\sum_{l=1}^{m+n} t_l = 1$, $t_l \geq 0$. Thus, it is in the convex hull, and so $\mathrm{Conv}(A)$ is convex.

Let

$$\mathcal{F} = \{U : A \subset U, U \text{ is convex}\}.$$

We claim that

$$\operatorname{Conv}(A) = \bigcap_{U \in \mathcal{F}} U.$$

First, note that we've shown $\operatorname{Conv}(A)$ is convex, and it's clear that $A \subset \operatorname{Conv}(A)$ by choosing $n = 1$ and $t = 1$. Hence, $\operatorname{Conv}(A) \in \mathcal{F}$, so

$$\bigcap_{U \in \mathcal{F}} U \subset \operatorname{Conv}(A).$$

So, we need to show that if V is convex and

$$A \subset V \subset \operatorname{Conv}(A),$$

then $V = \operatorname{Conv}(A)$ (in other words, $\operatorname{Conv}(A)$ is the smallest convex set containing A). Let $x \in \operatorname{Conv}(A)$, then

$$x = \sum_{i=1}^{n} t_i a_i,$$

where $\sum_{i=1}^{n} t_i = 1$, $t_i \geq 0$, $a_i \in A$. We inductively show that this is in V. In the case $n = 1$, it's clear (since $A \subset V$). In the case $n = 2$, we have

$$x = ta_1 + (1-t)a_2 \in tA + (1-t)A \subset tV + (1-t)V \subset V.$$

Hence, $x \in V$. Assume that this construction works up to $n - 1$. We need to show that x in the form above is in V. We write

$$\sum_{i=1}^{n} t_i a_i = \sum_{i=1}^{n-1} t_i a_i + t_n a_n.$$

Notice that $\sum_{i=1}^{n-1} = 1 - t_n$. If $t_n = 1$, this is 0 and it's clear that this is going to be in V, so assume $t_n < 1$. If $t_n = 0$, we have that this is in the form $\sum_{i=1}^{n-1}$, and so the induction hypothesis says this is in V and we win. So assume $0 < t_n < 1$, so that $0 < \sum_{i=1}^{n-1} t_i < 1$. We can then normalize to get

$$\frac{\sum_{i=1}^{n-1} t_i a_i}{\sum_{i=1}^{n-1} t_i} = z \in V$$

by the induction hypothesis, so

$$\sum_{i=1}^{n-1} t_i a_i = \left(\sum_{i=1}^{n-1} t_i \right) z.$$

Hence,

$$x = \sum_{i=1}^{n} t_i a_i = \sum_{i=1}^{n-1} t_i a_i + t_n a_n = \left(\sum_{i=1}^{n-1} t_i \right) z + t_n a_n = (1 - t_n) z + t_n a_n \in (1 - t_n)V + t_n V \subset V,$$

so it holds up to n. Hence, induction applies, and so we get that all $x \in \operatorname{Conv}(A)$ are such that $x \in V$, so $V = \operatorname{Conv}(A)$. So for all U convex with $A \subset U$, we have $\operatorname{Conv}(A) \subset U$, and hence $\bigcap_{U \in \mathcal{F}} U = \operatorname{Conv}(A)$. □

Remark. This gives us two ways to think about the convex hull – a constructive way (convex combinations of elements) and a theoretical way (intersection of all convex sets containing A).

Problem 3 (Rudin 1.3, James)**.** Let X be a topological vector space. All sets mentioned below are understood to be the subsets of X. Prove the following statements.

(1) The convex hull of every open set is open.
(2) If X is locally convex, then the convex hull of every bounded set is bounded.
(3) If A and B are bounded, so is $A + B$.

(4) If A and B are compact, so is $A+B$.
(5) If A is compact and B is closed, then $A+B$ is closed.
(6) The sum of two closed sets may fail to be closed.

Proof.

(1) See **Problem 2, part (3)**.
(2) Let $E \subset X$ be a bounded set. Then for every neighborhood U of 0, there exists a t such that for all $s > t$, we have $E \subset sU$. In particular, we can choose this U to be convex (since X is locally convex). We claim that sU is convex. This follows, since
$$tsU + (1-t)sU = s(tU) + s((1-t)U) = s\left[tU + (1-t)U\right] \subset sU.$$
Hence, $\text{Conv}(E) \subset sU$ for all $s > t$. This implies that $\text{Conv}(E)$ is convex.
(3) Let V be a neighborhood of 0. Since A and B are both bounded, there exists t_1 and t_2 so that $A \subset sV$ for all $s > t_1$ and $B \subset sV$ for all $s > t_2$. In particular, choosing $t = \max\{t_1, t_2\}$, we have that $A + B \subset sV + sV = (s+s)V$ for $s > t$. Hence, choosing $2t$, we have that for all $s > 2t$, $A + B \subset sV$. We can do this for all neighborhoods of 0, so $A+B$ is bounded.
(4) Addition is continuous, and $A+B$ is the image of a compact set $A \times B$.
(5) The goal is to show that if $x \notin A+B$, then there is a neighborhood of x which does not intersect B. We can write
$$A + B = \bigcup_{a \in A} a + B,$$
so $x \notin A + B$ implies $x \notin a + B$ for all $a \in A$. Since B is closed, $a + B$ is also closed for all $a \in A$. So we have $\{x\} \cap (a+B) = \varnothing$, and so we can find neighborhood U_a and V_a with $x \in U_a$ and $(a+B) \subset V_a$ with $U_a \cap V_a = \varnothing$. Now, note that
$$V_a - B = \bigcup_{b \in B} (V_a - b)$$
is an open set which contains a, so
$$A \subset \bigcup_{a \in A} (V_a - B),$$
and since A is compact, we have
$$A \subset \bigcup_{j=1}^{n} (V_j - B)$$
for some finite refinement. Choose U_j which corresponds to V_j, then let $U = \bigcap U_j$. This is a finite intersection of open sets, so open, and $x \in U$. We have that $U \cap (A+B) = \varnothing$, since otherwise $y = a + b \in U \cap (A+B)$, so $y \in V_j$ for some j and $y \in U_j$, which is a contradiction.
(6) Consider $X = \mathbb{R}$, $A = \{n : n \geq 1\} = \bigcup_{n \geq 1} \{n\}$. Singletons are closed, so this is a closed set. Now, consider $B = \{-n + 1/n : n \geq 1\}$. This is closed, since $\lim_{n\to\infty}(1/n - n) = -\infty$. We have $A + B = \{1/n : n \geq 1\}$, which is not closed since $0 \notin A + B$.

□

Problem 4 (Rudin 1., James)**.** Let $B = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| \leq |z_2|\}$. Show that B is balanced, but its interior is not.

Proof. Let $\alpha \in \mathbb{C}$ be such that $|\alpha| \leq 1$. Then
$$\alpha B = \{(\alpha z_1, \alpha z_2) \in \mathbb{C}^2 : |z_1| \leq |z_2|\}.$$
Notice that $|\alpha z_1| = |\alpha||z_1| \leq |\alpha||z_2| = |\alpha z_2|$, so $\alpha B \subset B$. Hence, it is balanced.

The interior is $B^{\circ} = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| < |z_2|\}$, and it is not balanced since $0B^{\circ} = 0 \notin B^{\circ}$. □

Problem 5 (Rudin 1.5, James)**.** Consider the definition of "bounded set" given in Section 1.6. Would the content of this definition be altered if it were required merely that to every neighborhood V of 0 corresponds some $t > 0$ such that $E \subset tV$?

Proof. Recall that a subset E of a topological vector space is said to be *bounded* if to every neighborhood V of 0 in X corresponds a number $s > 0$ such that $E \subset tV$ for every $t > s$. Call this condition (1).

As mentioned on Wikipedia, the definitions are in fact equivalent. Labeling the condition in the problem condition (2), the goal is to show that (2) $\implies$ (1). Let E be a set which satisfies (2) so that for every neighborhood V of 0, we have that there is a $t > 0$ so that $E \subset tV$. Using **Rudin Theorem 1.14**, we can assume without loss of generality that V is a balanced neighborhood of 0. Consider $s \geq t$. The goal is to show that $E \subset sV$ for all such s. Since $s > t$, we have that $1 > t/s$, so since V is balanced we have that $tV = (t/s)sV = s((t/s)V) \subset sV$. Hence $E \subset tV \subset sV$ for all $s \geq t$, so we have that there exists a number $t > 0$ so that $E \subset tV$ for all $s > t$. E is then a balanced set with respect to condition (1). □

Problem 6 (Rudin 1.6, James)**.** Prove that a topological vector space is bounded if and only if every countable subset of E is bounded.

Proof. If $E \subset X$ is bounded and $F \subset E$ is countable, then for every neighborhood of the origin V we have that there exists a $t > 0$ so that $F \subset E \subset tV$, and hence F is bounded.

Assume now that E is not bounded. There exists a neighborhood of the origin V so that for all $t > 0$, $E \not\subset tV$. So for every $n \geq 1$, we see that $E \not\subset nV$, so we can choose $x_n \in E \cap (nV)^c$. The set $\{x_n\} \subset E$ is countable, and as seen in the last problem this implies that $\{x_n\} \not\subset tV$ for all $t > 0$. □

Problem 7 (Rudin 1.7, James)**.** Let

$$X = \{f : [0,1] \to \mathbb{C}\}.$$

Topologize X by the family of seminorms

$$\rho_x(f) = |f(x)| \qquad (0 \leq x \leq 1).$$

The topology is called the *topology of pointwise convergence.*

(1) Justify this terminology.
(2) Show that there exists a sequence $\{f_n\} \subset X$ such that $\{f_n\}$ converges to 0 as $n \to \infty$, but if $\{\gamma_n\}$ is any sequence of scalars such that $\gamma_n \to \infty$, then $\{\gamma_n f_n\}$ does not converge to 0.
(3) Deduce that metrizability cannot be omitted in (b) of Theorem 1.28.

Proof.

(1) Suppose $f_n \to f$ with respect to this topology, then we have that $\rho_x(f_n - f) \to 0$. This implies that $|f_n(x) - f(x)| \to 0$ for all $x \in [0,1]$. Hence $f_n \to f$ pointwise.
(2) As Thomas suggested, the space of all complex sequences which do not contain 0 and converge to 0 are in bijection with $[0,1]$, so for any sequence γ_n such that $\gamma_n \to 0$ and does not contain 0 we can associate it to a point x. Every sequence $\delta_n \to \infty$ can be written as γ_n^{-1} for some sequence $\gamma_n \to 0$ which does not contain 0. For each $x \in [0,1]$, let $f_n(x) = \gamma_n$, and consider the sequence γ_n^{-1}. For each point x, we have $f_n(x)\gamma_n^{-1} \to 1$ for the associated γ_n, and hence $f_n\gamma_n^{-1} \not\to 0$.
(3) **Rudin Theorem 1.28** says that if X is metrizable, then for any sequence $\{x_n\}$ which converges to 0, there exists a sequence $\gamma_n \to \infty$ with $\gamma_n x_n \to 0$. (2) contradicts this property.

□

Problem 8 (Rudin 1.8, James)**.**

(1) Suppose $\mathcal{F}$ is a separating family of seminorms on a vector space X. Let $\mathcal{L}$ be the smallest family of seminorms on X that contains $\mathcal{F}$ and is closed under max. If the construction of **Rudin Theorem 1.37** is applied to $\mathcal{F}$ and $\mathcal{L}$, show that that the two resulting topologies coincide. The main difference is that $\mathcal{L}$ leads directly to a base, rather than a subbase.

(2) Suppose $\mathcal{L}$ is as in part (1) and Λ is a linear functional on X. Show that Λ is continuous if and only if there exists a $p \in \mathcal{L}$ such that

$$|\Lambda x| \leq Mp(x)$$

for all $x \in X$ and some constant $M < \infty$.

Proof. The topology from **Rudin Theorem 1.37** is constructed by taking finite intersections of sets of the form

$$V(p,n) = \left\{ x \in X : p(x) < \frac{1}{n} \right\}$$

for all $p \in \mathcal{F}$ and then taking translates. The local base $\mathcal{B}$ generated by the finite intersections of these sets is what we're really concerned with.

(1) Note that if τ is the topology generated by $\mathcal{F}$ and θ is the topology generated by $\mathcal{L}$, then since $\mathcal{F} \subset \mathcal{L}$ we have that $V(p,n) \in \theta$ for all $p \in \mathcal{F}$, so that $\tau \subset \theta$. We then need to show that $\theta \subset \tau$. To do so, it suffices to show that if $\mathcal{B}_1$ is the local base associated to $\mathcal{F}$, $\mathcal{B}_2$ is the local base associated to $\mathcal{L}$, then for every $U \in \mathcal{B}_2$ there is a $V \in \mathcal{B}_1$ with $V \subset U$.

The idea is to note that if $p = \max\{p_1, p_2\}$, then we have that

$$V(p,n) = V(p_1,n) \cap V(p_2,n).$$

Once we have this, the result follows. This is also easy, since $x \in X$ satisfies $p(x) < 1/n$ if and only if $p_1(x) < 1/n$ and $p_2(x) < 1/n$. This then gives us that the local bases generated are the same, so the topologies generated are the same. It also follows that it leads to a base, rather than a subbase, since we don't need to concern ourselves with finite intersections.

(2) ($\Longrightarrow$): Assume that Λ is continuous. Being continuous is equivalent to being continuous at the origin. We have that for the open ball $B_1(0) \subset \mathbb{R}$, $\Lambda^{-1}(B_1(0)) \subset X$ is open. Notice

$$\Lambda^{-1}(B_1(0)) = \{x \in X : |\Lambda(x)| < 1\}.$$

Since $V(p,n)$ forms a base for our topology, we can find p and n so that

$$V(p,n) \subset \Lambda^{-1}(B_1(0)).$$

So $x \in X$ with $p(x) < 1/n$ implies $|\Lambda(x)| < 1$, so we have $|\Lambda(x)| \leq np(x)$ for all $x \in X$.
($\Longleftarrow$): The assumption tells us that $V(p,M) \subset \Lambda^{-1}(B_1(0))$. Notice that for any $\epsilon > 0$, we have $\epsilon V(p,M) \subset \epsilon\Lambda^{-1}(B_1(0))$. Notice that

$$\begin{aligned}\epsilon\Lambda^{-1}(B_1(0)) = \{\epsilon x : x \in X, |\Lambda(x)| < 1\} &= \{x \in X : |\Lambda(\epsilon^{-1}x)\}| < 1\} \\ &= \{x \in X : |\Lambda(x)| < \epsilon\} = \Lambda^{-1}(B_\epsilon(0)).\end{aligned}$$

So for every ball $B_\epsilon(0)$, we have $\epsilon V(p,M) \subset \Lambda^{-1}(B_\epsilon(0))$. Notice $\epsilon V(p,M)$ is an open set in θ so Λ is continuous at the origin. **Rudin Theorem 1.17** tells us that Λ is continuous.

□

Problem 9 (Rudin 1.9, James)**.** Suppose

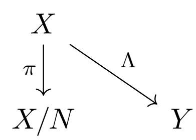
(i) X and Y are topological vector space,
(ii) $\Lambda : X \to Y$ is linear,
(iii) N is a closed subspace of X,

(iv) $\pi : X \to X/N$ is the quotient map, and
(v) $\Lambda x = 0$ for every $x \in N$.

(1) Prove that there is a unique $f : X/N \to Y$ which satisfies $\Lambda = f \circ \pi$.
(2) Prove that this f is linear.
(3) Prove that Λ is continuous if and only if f is continuous.
(4) Prove that Λ is open if and only if f is open.

Proof. Throughout, we utilize **Rudin Theorem 1.41** which establishes that $\pi : X \to X/N$ is an open, continuous, linear map.

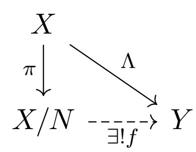
(1) In a diagram, we have the following:



We define the map $f : X/N \to Y$ via $f(x + N) = \Lambda(x)$. We first check this is well-defined. If $x - y \in N$, then we have that

$$f(x + N) = \Lambda(x) = \Lambda(x + y - y) = \Lambda(x - y) + \Lambda(y) = \Lambda(y) = f(y + N).$$

The map is therefore well-defined. Notice it is such that $f \circ \pi = \Lambda$, since $f \circ \pi(x) = f(x + N) = \Lambda(x)$. We check that it is unique. If g is another function so that $g \circ \pi = \Lambda$, then for each $x \in X$ we have that $g(x + N) = \Lambda(x) = f(x + N)$. Every element in X/N is in this form, so we have that $f = g$ as functions. In a diagram, we have



commutes.

(2) We now check that the map is linear. Let α be a scalar, $x \in X$, then we have that

$$f(\alpha(x + N)) = f(\alpha x + N) = \Lambda(\alpha x) = \alpha\Lambda(x) = \alpha f(x + N),$$

and if $x, y \in N$, then we have

$$f((x + N) + (y + N)) = f((x + y) + N) = \Lambda(x + y) = \Lambda(x) + \Lambda(y) = f(x + N) + f(y + N),$$

where in both cases we utilized the fact that N is a (closed) subspace.

(3) Let $U \subset Y$ be open. If f is continuous, we have that $f^{-1}(U)$ is open, and by the continuity of π we have $\pi^{-1}(f^{-1}(U)) = \Lambda^{-1}(U)$ is open. This applies for all open sets, so Λ is continuous.

If Λ is continuous, then for all $U \subset Y$ open we have $\Lambda^{-1}(U) = \pi^{-1}(f^{-1}(U))$ is open. We note that π is an open map, so applying π we get that $\pi(\pi^{-1}(f^{-1}(U))) = f^{-1}(U)$ is an open set. This applies for all $U \subset Y$ open, so we have f is continuous.

(4) Let $U \subset X$ be open. If Λ is open, we have $\Lambda(U) = f(\pi(U))$ is open. Notice that every open subset of X/N is realized by $\pi(U)$ for some open subset $U \subset X$, so we get that for every open subset $V \subset X/N$, $f(V)$ is open.

If f is an open map, then for all $V \subset X/N$ open we have that $f(V)$ is open. Let $U \subset X$ be open, then $\Lambda(U) = f(\pi(U))$ is open, since π is an open map. We get that Λ is an open map.

□

Problem 10 (Rudin 1.10, James)**.** Suppose X and Y are topological vector spaces, $\dim(Y) < \infty$, $\Lambda : X \to Y$ is linear, and $\Lambda(X) = Y$ (so that Λ is surjective).

(1) Prove that Λ is an open mapping.
(2) Assume, in addition, that the null space of Λ is closed. Prove that Λ is continuous.

Proof. Let $N = \ker(\Lambda)$, then $X/N \cong Y$ by surjectivity and the first isomorphism theorem. Suppose N is closed. By the last problem there exists a map $f : X/N \to Y$ so that $\Lambda = f \circ \pi$. This map is linear, and if N is closed then by the last problem and **Rudin Theorem 1.21**, we have that the map is a homeomorphism.

If N is not closed, we will have that X/N is not a topological vector space. We have to put in more work for this case.

Let $\{e_i\}$ be a basis for Y. Let $a_i \in \Lambda^{-1}(e_i)$. Since the e_i are linearly independent, we have that the a_i are also linearly independent, since $\sum t_i a_i = 0$ implies $\Lambda(\sum t_i a_i) = \sum t_i \Lambda(a_i) = \sum t_i e_i = 0$, so $t_i = 0$ for all i. The map $F : \mathbb{R}^n \to X$ given by $F(x_1, \ldots, x_n) = \sum x_i a_i$ is a linear map. Invoking **Rudin Lemma 1.20**, it is continuous, and we note that $\ker(F) = 0$ by linear independence.

Let $\Gamma : \mathbb{R}^n \to Y$ be defined by $\Gamma = \Lambda \circ F$. By the surjectivity of Λ and injectivity of F, we get Γ is bijective. Let $U \subset X$ be a neighborhood of the origin. F continuous implies $F^{-1}(U)$ is open and we note it contains the origin, so there is a $\epsilon > 0$ with $B_\epsilon(0) \subset F^{-1}(U)$. Γ is a bijective linear map between finite dimensional vector spaces so a homeomorphism. Hence, $\Gamma(B_\epsilon(0))$ is an open subset containing the origin, and we see that $\Gamma(B_\epsilon(0)) \subset \Lambda(U)$, so that the image of every open neighborhood of the origin in X contains an open neighborhood of the origin in Y. □

Remark. Question: Did we really need to do all of that work for the second part? This (among other resources) suggests yes, but I want to use **Rudin Exercise 1.9** to do it faster.

1.2. Chapter 2.

Problem 11 (Rudin 2.1, James)**.**

(1) If X is an infinite-dimensional topological vector space which is the union of countably many finite-dimensional subspaces, prove that X is of the first category in itself.
(2) Prove that no infinite-dimensional F-space has a countable Hamel basis.

Proof.

(1) First, we remark that a finite-dimensional subspace $V \subset X$ is a closed subspace. (**Rudin Theorem 1.21 (b)**). Next, we claim that a closed subspace of X must be meager

Claim. Let $V \subset X$ be a closed subspace. If V has non-empty interior then $V = X$.

Proof. Suppose that $E \subset V$ is a nonempty open subset. Let $a \in E$. Then $E - a$ is an open neighborhood of the origin. Since V is a subspace, we have that $E - a \subset V$ still. Note that for arbitrary $x \in X$, we have that there exists n sufficiently large so that $(1/n)x \in E - a \subset V$. A vector subspace is closed under scalar multiplication, so $n((1/n)x) = x \in V$. The choice of x was arbitrary, giving us $X = V$. □

Taking the contrapositive of the above claim, we have that $V \neq X$ implies that V is meager. Thus, we have that

$$X = \bigcup_{n=1}^{\infty} V_n,$$

where the $V_n \subset X$ are finite-dimensional subspaces. This gives us that X is a countable union of meager sets, and so it is of first category in itself.

(2) An F-space is a complete metric space, so Baire's theorem (**Rudin Theorem 2.2**, see discussion below) tells us that X is of second category. If X admits a Hamel basis β, then every $x \in X$ admits a unique representation of the form

$$x = \sum_{i=1}^{n} \alpha_i v_i,$$

where $\{v_i\}_{i=1}^n \subset \beta$.

Suppose β is countable. For every finite subset $\alpha \subset \beta$, we can associate a finite dimensional vector subspace V_α by taking the span of the vectors in α. Since β is a Hamel basis, the above discussion tells us that

$$X = \bigcup_{\alpha \subset \beta} V_\alpha.$$

This is a countable union of finite-dimensional vector subspaces, so (1) tells us that X is of first category, contradicting the fact that X is a F-space.

□

Problem 12 (Rudin 2.2, James)**.** Sets of first and second category are "small" and "large" in a topological sense. These notions are different when "small" and "large" are understood in the sense of measure, even when the measure is intimately related to the topology. To see this, construct a subset of the unit interval which is of the first category but whose Lebesgue measure is 1.

Proof. Take a fat Cantor set. □

Problem 13 (Rudin 2.3, James)**.** Put $K = [-1, 1]$. Define

$$\mathcal{D}_K = \{f \in C^\infty(\mathbb{R}) : \operatorname{supp}(f) \subset K\}.$$

Let $\{f_n\} \subset L^1$ be such that

$$\Lambda\varphi = \lim_{n\to\infty} \int_K f_n(t)\varphi(t)dt$$

exists for every $\varphi \in \mathcal{D}_K$.

(1) Show that Λ is a continuous linear functional on $\mathcal{D}_K$.

(2) Let $f_n(t) = n^3 t\chi_{[-1/n,1/n]}$. Show that there exists an $M < \infty$ so that

$$\left|\int_K f_n(t)\varphi(t)dt\right| \le M\|D\varphi\|_\infty$$

for all n.

(3) Construct an example where there exists $M < \infty$ so that

$$\left|\int_K f_n(t)\varphi(t)dt\right| \le M\|D^2\varphi\|_\infty$$

for all n, but there is no such M so that

$$\left|\int_K f_n(t)\varphi(t)dt\right| \le M\|D\varphi\|_\infty$$

for all n.

(4) Show that in general there exists a positive integer p and a number $M < \infty$ so that

$$\left|\int_K f_n(t)\varphi(t)dt\right| \le M\|D^p\varphi\|_\infty.$$

Proof.

(1) Define

$$\Lambda_n = \int_K f_n(t)\varphi(t).$$

The linearity of the integral tells us that $\{\Lambda_n\}$ is a family of linear functions from $\mathcal{D}_K \to \mathbb{R}$. We note that these are continuous utilizing the dominated convergence theorem. By the discussion in **Rudin Section 1.46**, we note that $\mathcal{D}_K$ is a Frechet space, so in particular an F-space. By **Rudin Theorem 2.8**, we get that Λ is continuous.

(2) Note that $f_n(t) \in L^1$ for all n, since

$$\int_K f_n(t)dt = \int_{-1/n}^{1/n} n^3 t dt = n^3 \left(\frac{t^2}{2}\Big|_{-1/n}^{1/n} \right) = \frac{n^3}{t}\left(\frac{1}{n^2} - \frac{1}{n^2}\right) = 0.$$

Note that for any constant C we have

$$F_n(t) = n^3 \frac{t^2}{2} + C$$

is such that $F_n' = f_n$. Without loss of generality, we choose $C = 0$.

Let $\varphi \in \mathcal{D}_K$. Using integration by parts, we have

$$\int_K f_n(t)\varphi(t)dt = \int_{-1/n}^{1/n} f_n(t)\varphi(t)dt = \frac{n^3 t^2}{2}\varphi(t)\Big|_{-1/n}^{1/n} - \frac{n^3}{2}\int_{-1/n}^{1/n} D\varphi(t)t^2 dt.$$

Notice

$$\frac{n^3 t^2}{2}\varphi(t)\Big|_{-1/n}^{1/n} = \frac{n}{2}\varphi(1/n) - \frac{n}{2}\varphi(-1/n) = \frac{n}{2}\left[\varphi(1/n) - \varphi(-1/n)\right] = \frac{\varphi(1/n) - \varphi(-1/n)}{2/n}.$$

By the mean value theorem, we have that this is bounded above by $\|D\varphi\|_\infty$. We similarly note that

$$-\frac{n^3}{2}\int_{-1/n}^{1/n} D\varphi(t)t^2 dt \leq -\|D\varphi\|_\infty \frac{n^3}{2}\int_{-1/n}^{1/n} t^2 dt = -\frac{1}{3}\|D\varphi\|_\infty.$$

Hence we have that

$$\left|\int_K f_n(t)\varphi(t)dt\right| \leq \frac{2}{3}\|D\varphi\|_\infty.$$

(3) TODO

(4) This follows by **Rudin Exercise 1.8** and the fact that the topology is generated by the seminorms

$$p_N(f) = \max\{|D^\alpha f(x)| : x \in K_N, |\alpha| \leq N\}.$$

□

Problem 14 (Rudin 2.4, James)**.** Let L^1 and L^2 be the usual Lebesgue spaces on the unit interval. Prove that $L^2 \subset L^1$ is of the first category in three ways:

(1) Show that

$$A_n = \left\{ f : \int |f|^2 \leq n \right\}$$

is closed in L^1 but has empty interior.

(2) Put $g_n = n$ on $[0, n^{-3}]$ and show that

$$\int f g_n \to 0$$

for every $f \in L^2$ but not every $f \in L^1$.

(3) Show that the inclusion map of L^2 into L^1 is continuous, but not onto.

Proof.

(1) We establish the set is closed. Let $\{f_m\} \subset A_n$ be a sequence with $f_m \to f$ in L^1. The goal is to show that $f \in A_n$ as well. Recall $f_m \to f$ in L^1 if

$$\int |f - f_m| \to 0.$$

Convergence in L^1 implies convergence in measure, and since $f_m \to f$ in measure we have a subsequence $f_{m_j} \to f$ almost everywhere. Notice the sequence $|f_{m_j}|^2 \to |f|^2$ almost everywhere as well. Finally for each m_j we have

$$\int |f_{m_j}|^2 \leq n.$$

Applying Fatou's lemma gives us

$$\int |f|^2 \leq \liminf_{j\to\infty} \int |f_{m_j}|^2 \leq n.$$

So $f \in A_n$.

Next, we need to show that the interior of A_n is empty for all n. Let $f \in A_n$, and consider the ball

$$B_k(f) = \left\{ g \in L^1 : \int |f - g| < \frac{1}{k} \right\}.$$

The goal is to show that there is no k with $B_k(f) \subset A_n$. Let $g \in L^2 \setminus L^1$, and notice that for $r > \|g\|_1/k$, we have that $f + (1/r)g \in B_k(f)$. This follows from

$$\int |f - f - (1/r)g| = \frac{\|g\|_1}{r} < k.$$

We now claim that $f + (1/r)g \notin A_n$. If $f + (1/r)g$ were in A_n, we would have

$$\int |f + (1/r)g|^2 = \frac{1}{r^2} \int |rf + g|^2 \leq n.$$

In other words,

$$\int |rf + g|^2 < \infty,$$

so $rf + g \in L^2$. In other words, there is some $h \in L^2$ with $rf + g = h$. We note that L^2 is a vector subspace, so this implies that $g = h - rf \in L^2$. This contradicts the fact that $g \in L^1 \setminus L^2$, so we cannot have $f + (1/r)g \in A_n$. The choice of k was arbitrary, so for all k we get that

$$B_k(f) \not\subset A_n,$$

and so f is not in the interior of A_n. The choice of f was arbitrary, so $(A_n)^{\mathrm{o}} = \varnothing$.

This tells us that the A_n are meager sets, and we have that

$$L^2 = \bigcup_{n\geq 1} A_n,$$

so L^2 is of first category in L^1.

(2) Let $f \in L^2$. By Hölder's inequality, we have

$$\int f g_n = \int_0^{n^{-3}} nf \leq \|f\|_2 \|n\chi_{[0,n^{-3}]}\|_2,$$

and we note that

$$\left(\int_0^{n^{-3}} n^2 \right)^{1/2} = \frac{1}{n},$$

so as $n \to \infty$ we get that $\int f g_n \to 0$.

Note the same trick doesn't work here, since for $f \in L^1$ we have

$$\int f g_n \leq \|f\|_1 \|n\chi_{[0,n^{-3}]}\|_\infty = \|f\|_1 n \to \infty.$$

Let

$$f(x) = x^{-3/4}\chi_{[0,1]}.$$

Then

$$\|f\|_1 = 4 < \infty,$$

and we have

$$\int f g_n = n\int_0^{n^{-3}} x^{-3/4}dx = 4n^{1/4} \to \infty.$$

Define

$$\Lambda_n : L^1 \to \mathbb{R}, \qquad \Lambda_n(f) = \int f g_n.$$

By what we observed above, each Λ_n is continuous and linear. We have that the collection of all points whose orbits under the action of $\Gamma = \{\Lambda_n\}$ are bounded is going to be $B = L^2$. We showed that $B \neq L^1$, so by the contrapositive of Banach-Steinhaus we cannot have that B is of the second category.

(3) Consider the map $T : L^2 \to L^1$ which is the identity. Since we're on the unit interval, we have $L^2 \subset L^1$, so this is fine. To see that it's continuous, we recall that

$$\|T(f)\|_1 = \|f\|_1 \leq \|f\|_2.$$

This follows by using Hölder's inequality:

$$\|f\|_1 = \int |f| \cdot 1 \leq \||f|^1\|_{2/1} \cdot \|1\|_{2/(2-1)} = \|f\|_2 \mu([0,1])^{(2-1)/2} = \|f\|_2.$$

The inclusion is therefore continuous. Note that T is not surjective, since $\frac{1}{\sqrt{x}} \notin L^2$. The contrapositive of the open mapping theorem says that L^2 cannot be of second category in L^1.

□

Problem 15 (Rudin 2.6, James)**.** Define the Fourier coefficients $\widehat{f}(n)$ of a function $f \in L^2(T)$ (where T is the unit circle) by

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}d\theta$$

for all $n \in \mathbb{Z}$. Put

$$\Lambda_n f = \sum_{k=-n}^{n} \widehat{f}(k).$$

Prove that

$$L := \left\{ f \in L^2(T) : \lim_{n\to\infty} \Lambda_n f \text{ exists.} \right\}$$

is a dense subspace of $L^2(T)$ of the first category.

Remark. Folland defines Fourier coefficients by

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{-in\theta}d\theta.$$

This is fine, except I'm imagining the function is already a function on the circle so there's no need to compose with $e^{i\theta}$ (see Wikipedia for example).

Proof. We first show that A is dense. Stone-Weierstrass says that trigonometric polynomials are dense in $L^2(T)$, so if we show that the series of Fourier coefficients converges for trigonometric polynomials then we win. By linearity, it boils down to showing that if $f_k(x) = \exp(ikx)$, then $\lim_{n\to\infty} \Lambda_n f_k$ exists.

To see this, notice that

$$\widehat{f_k}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ikx}e^{-inx}dx = \begin{cases} 1 \text{ if } n = k \\ 0 \text{ otherwise.} \end{cases}$$

Hence we see that for sufficiently large n, we get

$$\Lambda_n f_k = \sum_{-n}^{n} \widehat{f_k}(j) = 1.$$

So $\lim_{n\to\infty} \Lambda_n f_k$ exists for this function, and so for linear combinations, and so for trigonometric polynomials.

Note that A is a subspace (imagining $L^2(T)$ as a vector space). This directly follows from the linearity of limits and the fact that Λ_n is a linear function with respect to n.

Let $\Gamma = \{\Lambda_n\}$. We have $\Lambda_n : L^2(T) \to \mathbb{R}$ is a linear functional on $L^2(T)$ for each n. Suppose that L is of the second category. **Rudin Theorem 2.7 (b)** says that, since $\mathbb{R}$ is a F-space, $L = X$ and $\Lambda : X \to Y$ is continuous, where Λ is the limit. However, we have that $L \neq X$, forcing it to be of the first category. To see this, consider

$$-if(x) = \begin{cases} -i\pi + ix \text{ if } 0 < x < \pi \\ i\pi + ix \text{ if } -\pi < x < 0 \end{cases}$$

(see here for how to construct such a function). This is a function where $\widehat{f}(n) = \frac{1}{n}$ for $n \neq 0$, and so the limit of $\Lambda_n f$ does not exist. Hence, $L \neq X$. □

Remark. This is weird, since it is dense but it is the union of nowhere dense sets.

Problem 16 (Rudin 2.7, James)**.** Let $C(\mathbb{T})$ be the set of all continuous complex functions on the unit circle $\mathbb{T}$. Suppose $\{\gamma_n\}$ is a complex sequence that associates to each $f \in C(\mathbb{T})$ a function $\Lambda f \in C(\mathbb{T})$ whose Fourier coefficients are

$$\widehat{\Lambda f}(n) = \gamma_n \widehat{f}(n).$$

Prove that $\{\gamma_n\}$ has this multiplier property iff there is a complex Borel measure μ on T such that

$$\gamma_n = \int e^{-in\theta}d\mu(\theta).$$

Proof. We follow the hint given in Rudin. Since $\mathbb{T}$ is compact, we have that the infinity norm

$$\|f\|_\infty = \sup\{|f(x)| : x \in \mathbb{T}\}$$

is a well-defined norm on $C(\mathbb{T})$. One can show that $C(\mathbb{T})$ is a Banach space with this norm.

We check that $\Lambda : C(\mathbb{T}) \to C(\mathbb{T})$ is a linear function. Let $f, g \in C(\mathbb{T})$. Then $\Lambda(f + g)$ is a function in $C(\mathbb{T})$ whose Fourier coefficients are of the form

$$(\Lambda(f + g))^\wedge(n) = \gamma_n(f + g)^\wedge(n).$$

The Fourier transform is linear,so

$$(\Lambda(f + g))^\wedge(n) = \gamma_n\widehat{f}(n) + \gamma_n\widehat{g}(n) = (\Lambda f)^\wedge(n) + (\Lambda g)^\wedge(n).$$

Since the Fourier coefficients are all equal, we have that the functions $\Lambda(f + g) = \Lambda(f) + \Lambda(g)$ are equal almost everywhere, and since they are continuous this implies that they are equal everywhere. The same argument with scaling tells us that Λ is a linear function.

Now we have that $\Lambda : C(\mathbb{T}) \to C(\mathbb{T})$ is a linear function between Banach spaces, and we wish to show that it is continuous. The closed graph theorem tells us that it suffices to check the graph of Λ is closed. Let $f_n \to f$ in $C(\mathbb{T})$, then we need to show that $\Lambda(f_n) \to \Lambda(f)$. Note that the Fourier transform is continuous, and so with this we have

$$[\Lambda(f)]^\wedge(m) = \lim_{n\to\infty} [\Lambda(f_n)]^\wedge(m) = \lim_{n\to\infty} \gamma_m (f_n)^\wedge(m) = \gamma_m \widehat{f}(m).$$

By the same argument earlier, we get that $\lim_{n\to\infty} \Lambda(f_n) = \Lambda(f)$. Thus Λ is continuous.

Compose Λ with the map $f \mapsto f(0)$, which is a continuous linear functional. This gives us a new linear functional

$$f \mapsto (\Lambda f)(0) = \sum_{-\infty}^{\infty} \gamma_n \widehat{f}(n).$$

Note we will look at the space of trigonometric functions on $\mathbb{T}$, which is dense in $C(\mathbb{T})$ by a consequence of Fejers theorem. Notice that on the base $\kappa_n = \exp(-2\pi i n x)$, we have $(\Lambda\kappa_n)(0) = \gamma_n$. Extending by linearity gives us that it is a positive linear functional, so Riesz-Representation says that there is a unique complex Borel measure μ on $\mathbb{T}$ with

$$(\Lambda\kappa_n)(0) = \gamma_n = \int e^{-2\pi i n x} d\mu(x).$$

□

Problem 17 (Rudin 2.13, James)**.** Suppose X is a topological vector space which is of the second category in itself. Let K be a closed, convex, absorbing subset of X. Prove that K contains a neighborhood of the origin.

Proof. First, we show that if K is absorbing, then so is $-K$. If K is absorbing, then for every $x \in X$ there is a t_x so that for $s > t_x$, $x \in sK$. For every $x \in X$, there is a unique $(-x)$, so for this $(-x)$ there is a $t_{(-x)}$ so that for $s > t_{(-x)}$, $-x \in sK$, which implies $x \in s(-K)$. Since we can do this for all $x \in X$, we get that $-K$ is absorbing.

Next, if K_1, K_2 are two absorbing sets, we claim that $K_1 \cap K_2$ is absorbing. Let $x \in X$. Then there is t_1, t_2 so that for $s > t_1$, $x \in sK_1$, and for $s > t_2$ we ahve that $x \in sK_2$. Let $t = \max\{t_1, t_2\}$. Then for $s > t$, we have that $s > t_1$ and $s > t_2$, so $x \in sK_1$ and $x \in sK_2$. In other words, $x \in s(K_1 \cap K_2)$. We can do this for all $x \in X$, so the set is absorbing.

Since K and $-K$ are both absorbing, we have

$$H = K \cap (-K)$$

is absorbing. We now claim that $H^\circ \neq \varnothing$. Since H is absorbing, we have that

$$X = \bigcup_{n=1}^{\infty} nH.$$

Since X is of second category in itself, we get that there must be some n so that nH is not nowhere dense, and since scaling by n is a homeomorphism we get that H is nowhere dense. In other words, H has nonempty interior. Note that H is a closed, convex set since K is closed and convex.

Now we claim that

$$2H = H + H = H - H.$$

The first inequality follows from **Rudin Exercise 1.1 (4)**. The second inequality follows from the fact that $-H = H$. Since H has nonempty interior, let $U \subset H$ be open. We get that

$$U - U \subset H - H = 2H.$$

This is a nonempty open neighborhood of the origin. Using that scaling is a homeomorphism, we get that scaling down by $(1/2)$ still gives us an open neighborhood of the origin contained in $H \subset K$. So K contains a neighborhood of the origin. □

Problem 18 (Rudin 2.14, James)**.**

(1) Suppose X and Y are topological vector space, $\{\Lambda_n\}$ is an equicontinuous sequence of linear mappings from X into Y, and C is the set of all x at which $\{\Lambda_n(x)\}$ is a Cauchy sequence in Y. Prove that C is a closed subspace of X.

(2) In addition to the hypotheses of (1), assume that Y is an F space and that $\{\Lambda_n(x)\}$ converges in some dense subset of X. Prove that then

$$\Lambda(x) = \lim_{n\to\infty} \Lambda_n(x)$$

exists for every $x \in X$ and that Λ is continuous.

Proof.

(1) Let $x \in \overline{C}$. The goal is to show that $x \in C$. In other words, the goal is to show that $\{\Lambda_n(x)\}$ is a Cauchy sequence, meaning that for every open neighborhood of the origin V, we have that there is an N so that for $m, n \geq N$, we have

$$\Lambda_n(x) - \Lambda_m(x) \in V.$$

Let V be a neighborhood of the origin, and a balanced neighborhood of the origin U so that $U + U + U \subset V$. Since we have equicontinuity, choose a neighborhood W of the origin so that $\Lambda_n(W) \subset U$ for each W. Since $x \in \overline{C}$, we get that $(x + W) \cap C \neq \varnothing$, so there is a y in this intersection so that $\{\Lambda_n(y)\}$ is Cauchy. Now we have an N so that for $n, m \geq N$

$$\Lambda_n(y) - \Lambda_m(y) = \Lambda_n(y - x) + \Lambda_n(x) - \Lambda_m(y - x) + \Lambda_m(x) \in U$$

So

$$\Lambda_n(x) - \Lambda_m(x) = [\Lambda_n(y) - \Lambda_m(y)] - [\Lambda_n(y - x) - \Lambda_m(y - x)] \in U + U + U \subset V$$

(here using the fact that $y - x \in W$) and so the sequence is Cauchy. To see that it is a subspace is easy; if $x, y \in C$, then $x + y \in C$ by linearity. Same with scaling.

(2) By (1), the space of points where this converges is closed subspace. Since it contains a dense subset, it must be the whole space, so the limit exists for all $x \in X$. Hence, we can define

$$\Lambda(x) = \lim_{n\to\infty} \Lambda_n(x)$$

and this limit is well-defined. We check that this is continuous. Let $x, y \in X$, then

$$\Lambda(x + y) = \lim_{n\to\infty} \Lambda_n(x + y) = \lim_{n\to\infty} [\Lambda_n(x) + \Lambda_n(y)] = \Lambda(x) + \Lambda(y)$$

using linearity of the Λ_n and the limit. The same argument gives us scaling, so it is a linear map. For continuity, we show continuity at the origin and apply **Rudin Theorem 1.17**. Let V be a neighborhood of the origin in Y. By equicontinuity, there exists a neighborhood W of the origin in X so that $\Lambda_n(W) \subset V$ for all n. This gives us that $\Lambda(W) \subset V$, which tells us that Λ is continuous at the origin.

□

Remark. I filled in the details from here.

Problem 19 (Rudin 2.15, James)**.** Suppose X is an F-space and Y is a subspace of X whose complement is of the first category. Prove that $Y = X$.

Proof. The complement of Y is of the first category, so

$$Y^c = \bigcup_{\alpha=1}^{\infty} U_\alpha,$$

where the U_α are nowhere dense. Taking complements, we have that

$$Y = \bigcap_{\alpha=1}^{\infty} U_\alpha^c.$$

Taking individual U_α, we note that

$$\overline{U_\alpha}^{\mathrm{o}} = \varnothing,$$

so taking complements we have

$$\overline{(U_\alpha^c)^{\mathrm{o}}} = X;$$

that is, U_α^c contains an open set $V_\alpha \subset U_\alpha^c$ which is dense. So we have

$$\bigcap_{\alpha=1}^{\infty} V_\alpha \subset Y.$$

By **Baire's theorem** (**Rudin Theorem 2.2**) we get that the intersection of a countable collection of open dense sets is dense, so we get that Y must be dense.

Since Y is dense, we get that $Y \cap (Y + x) \neq \varnothing$, so take $y \in Y \cap (Y + x)$. Then $y - x \in Y$ so that $y - x = y'$ for some $y' \in Y$, hence $x = y - y' \in Y$. This holds for all $x \in X$, so $X = Y$. □

Problem 20 (Rudin 2.16, James)**.** Suppose that X and K are metric spaces, that K is compact, and that the graph of $f : X \to K$ is a closed subset of $X \times K$.

(1) Prove that f is continuous.
(2) Show that the compactness of K cannot be omitted from the hypotheses.

Proof.

(1) Consider

$$\Gamma(f) = \{(x, f(x)) : x \in X\} \subset X \times K.$$

Suppose $x_n \to x$. The goal is to show that $f(x_n) \to f(x)$. Suppose for contradiction that $f(x_n) \not\to f(x)$. Then we can construct a sequence $f(x_{n_j}) \to y \neq f(x)$ by compactness and the fact that if every convergent subsequence converges to $f(x)$, we have that the limit is $f(x)$. Then the sequence $(x_{n_j}, f(x_{n_j})) \subset \Gamma(f)$ is a sequence which converges to (x, y), but this then contradicts the fact that $\Gamma(f)$ is closed. Hence, we must have that every convergent subsequence converges to $f(x)$.

(2) Let $f : [0, 1] \to \mathbb{R}$ be $f(x) = 1/x$ if $x \neq 0$ and $f(0) = 0$. Then X is compact, and the graph of f is closed, but f is not continuous.

□

1.3. Chapter 3.

Problem 21 (Rudin 3.1, James)**.** Call a set $H \subset \mathbb{R}^n$ a hyperplane if there exists real numbers $a_1, \ldots, a_n, c$ (with $a_i \neq 0$ for at least one i) such that H consists of all of the points $x \in \mathbb{R}^n$ satisfying

$$\sum a_i x_i = c.$$

Suppose $E \subset \mathbb{R}^n$ is convex with nonempty interior, and $y \in \partial E$. Prove that there is a hyperplane H such that $y \in H$ and E lies entirely on one side of H.

Proof. Recall that the *Minkowski functional* μ_A of a subset A is defined by

$$\mu_A(x) = \inf\{t > 0 : t^{-1}x \in A\}.$$

This is well-defined and finite so long as A is absorbing.

Suppose 0 is an interior point of E. Since E has nonempty interior, this implies that the interior of E is a neighborhood of the origin. Hence, E is an absorbing set by prior claims. We can then define the Minkowski functional μ_E. As suggested in Rudin, the goal now is to try to apply **Theorem 3.2**. Notice that $\mu_E : X \to \mathbb{R}$ satisfies

$$\mu_E(x+y) \leq \mu_E(x) + \mu_E(y), \qquad \mu_E(tx) = t\mu_E(x)$$

by **Theorem 1.35**. So (b) of **Theorem 3.2** is satisfied.

Consider the subspace

$$M = \{\lambda y : \lambda \in \mathbb{R}\} \subset \mathbb{R}^n.$$

In other words, M is the subspace generated by y on the boundary of E. Define $f : M \to \mathbb{R}$ by

$$f(ty) = t.$$

This is a linear function, and we see that on M

$$f(ty) = t \leq \mu_E(ty).$$

Theorem 3.2 applies to give us $\Lambda : X \to \mathbb{R}$ with

$$\Lambda|_M = f, \qquad -\mu_E(-x) \leq \Lambda(x) \leq \mu_E(x).$$

Notice

$$\mu_E(x) \leq 1 \text{ for } x \in E \implies |\Lambda(x)| \leq 1 \text{ for } x \in E.$$

Notice $\Lambda(y) = 1$ as well.

Define H by

$$H = \{x \in \mathbb{R}^n : \Lambda(x) = 1\}.$$

We check that this is a hyperplane. Writing $x \in H$ as

$$x = (x_1, \ldots, x_n) = \sum x_i e_i,$$

we have

$$\Lambda(x) = \sum x_i \Lambda(e_i) = 1.$$

So if we set the constants $a_i = \Lambda(e_i)$, we see that this gives us a hyperplane with $c = 1$ (we remark that at least one of the a_i must be non-zero, since $\Lambda(y) \neq 0$). To be on one side of the hyperplane means that for all $z \in E$, we have

$$\sum z_i a_i \leq 1,$$

which we see is satisfied.

We have then proven the statement under the assumption that 0 is in the interior of E. If 0 is not in the interior of E, we can take an element z in the interior and shift E by z; i.e., we examine the set $F = E - z$. This is a convex set (since the translation of a convex set is convex) and it has 0 in the interior, so we can construct a hyperplane H' so that F lies entirely on one side of H'. We can then translate this back by setting $H = H' + z$. Notice that translating the hyperplane still gives us a hyperplane, since $y \in H$ is of the form $y = x + z$, $x \in H$, and applying our linear functional to this element gives

$$\Lambda(y) = \Lambda(x) + \Lambda(z) = \sum x_i \Lambda(e_i) + C = 1.$$

Thus we set $a_i = \Lambda(e_i)$ and $c = 1 - C$, giving us a hyperplane. Notice that E is still contained entirely on one side of the hyperplane. □

Remark. The hint on here was useful.

Problem 22 (Rudin 3.2, James)**.** Let $L^2 := L^2([-1,1])$ with respect to Lebesgue measure. Let

$$E_\alpha = \{f \in C([-1,1]) : f(0) = \alpha\}.$$

(1) Show that each E_α is convex.
(2) Show that each E_α is dense in L^2.

(3) Conclude that if $\alpha \neq \beta$, then E_α and E_β are two disjoint convex sets which cannot be separated by any continuous linear functional.

Proof. (1) Let $f, g \in E_\alpha$, $0 \leq t \leq 1$. The goal is to show

$$tf + (1-t)g \in E_\alpha.$$

Evaluating at 0, we have

$$tf(0) + (1-t)g(0) = t\alpha + (1-t)\alpha = \alpha,$$

so $tf + (1-t)g \in E_\alpha$.

(2) Fix α. Let $f \in L^2$. The goal is to show that for all $\epsilon > 0$, there exists a $g \in E_\alpha$ so that

$$\|g - f\|_2 < \epsilon.$$

Lebesgue measure is Radon, so in particular we have $C([-1,1]) \subset L^2$ is dense (**Folland Proposition 7.9**). So for any $h \in C([-1,1])$, if we can find $g \in E_\alpha$ with

$$\|g - h\|_2 < \epsilon/2,$$

then we are done by the triangle inequality. Consider an open ball $(-\delta, \delta) \subset [-1,1]$, and define a function $g(x)$ to be $h(x)$ when $x \notin (-\delta, \delta)$, $g(x)$ is the straight line connecting $h(-\delta)$ to α on $(-\delta, 0]$, and it is the straight line connecting α to $h(\delta)$ on $[0, \delta)$. This is continuous by construction, and we can do this for all $\delta > 0$. Let $N = \max\{\alpha, \|h\|_u\}$. Then

$$\left(\int_{-1}^{1} |g - h|^2\right)^{1/2} \leq \left(\int_{-\delta}^{\delta} 4N^2\right)^{1/2} = 2\sqrt{2}\sqrt{\delta}N.$$

Notice that N doesn't depend on δ, so we can choose

$$\delta < \frac{\epsilon^2}{16N^2}.$$

Then

$$\|g - h\|_2 < \epsilon/2,$$

as desired.

(3) Take $\alpha \neq \beta$, $\Lambda : L^2 \to \mathbb{R}$ continuous and linear. It's clear that $E_\alpha \cap E_\beta = \varnothing$, since if $f \in E_\alpha \cap E_\beta \implies f(0) = \alpha$ and $f(0) = \beta$, contradicting well-definedness. As long as $\Lambda(f) \neq 0$ for some $f \in L^2$, we have Λ is surjective by linearity. So $\Lambda(E_\alpha) = \mathbb{R}$, $\Lambda(E_\beta) = \mathbb{R}$, so these sets cannot be separated by a continuous linear functional.

□

Problem 23 (Rudin 3.4, James)**.** Let l^∞ be the space of all real bounded functions x on the positive integers. So $x : \mathbb{Z}_{\geq 0} \to \mathbb{R}$ with

$$\|x\|_\infty = \max\{|x_i| : i \geq 0\} < \infty.$$

Let τ be the translation operator defined on l^∞ by

$$(\tau x)(n) = x(n+1).$$

Prove that there exists a linear functional Λ on l^∞ such that

(1) $\Lambda(\tau x) = \Lambda x$,

(2)

$$\liminf_{n\to\infty} x(n) \leq \Lambda(x) \leq \limsup_{n\to\infty} x(n)$$

for every $x \in l^\infty$.

Proof. We follow the suggestion. Define

$$\Lambda_n(x) = \frac{\sum_{k=1}^n x(k)}{n},$$
$$M = \left\{x \in l^\infty : \lim_{n\to\infty} \Lambda_n(x) \text{ exists}\right\},$$
$$p(x) = \limsup_{n\to\infty} \Lambda_n(x).$$

The goal is to apply **Theorem 3.2**, so we need to check all of the assumptions.

(1) Let $x, y \in l^\infty$, α a scalar. We need to show

$$\alpha x + y \in M.$$

Notice that for each n, we have

$$\Lambda_n(\alpha x + y) = \frac{1}{n}\sum_{k=1}^n (\alpha x + y)(k) = \alpha \frac{1}{n}\sum_{k=1}^n x(k) + \frac{1}{n}\sum_{k=1}^n y(k)$$
$$\alpha\Lambda_n(x) + \Lambda_n(y)$$

Taking limits, we get

$$\lim_{n\to\infty} \Lambda_n(\alpha x + y) = \alpha \lim_{n\to\infty} \Lambda_n(x) + \lim_{n\to\infty} \Lambda_n(y).$$

So if the limits for x and y exist, they exist for linear combinations. So M is a linear subspace.

(2) Next, we need to check that $p : X \to \mathbb{R}$ satisfies the desired properties. Notice that

$$\limsup_{n\to\infty} \Lambda_n(x+y) = \limsup_{n\to\infty} [\Lambda_n(x) + \Lambda_n(y)] \le \limsup_{n\to\infty} \Lambda_n(x) + \limsup_{n\to\infty} \Lambda_n(y).$$

So

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

Notice that for $t \ge 0$, we have

$$p(tx) = \limsup_{n\to\infty} \Lambda_n(tx) = \limsup_{n\to\infty} [t\Lambda_n(x)] = t \limsup_{n\to\infty} \Lambda_n(x) = tp(x).$$

(3) Let $f : M \to \mathbb{R}$ be defined by

$$f(x) = \lim_{n\to\infty} \Lambda_n(x).$$

Clearly we have $f(x) \le p(x)$ on M and f is linear.

Thus, we get a linear functional $\Lambda : X \to \mathbb{R}$ with $\Lambda|_M = f$ and

$$-p(-x) = \liminf_{n\to\infty} \Lambda_n(x) \le \Lambda(x) \le \limsup_{n\to\infty} \Lambda_n(x) = p(x).$$

We claim now that

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^n x(k) \le \limsup_{n\to\infty} x(n).$$

To see this, fix an n and fix $j \le n$. Then we have

$$\frac{1}{n}\sum_{k=1}^n x(k) = \frac{1}{n}\sum_{k=1}^j x(k) + \frac{1}{n}\sum_{k=j+1}^n x(k).$$

Now notice

$$\frac{1}{n}\sum_{k=1}^n x(k) \le \frac{1}{n}\sum_{k=1}^j x(k) + \frac{n-j}{n}\sup_{p\ge j} x(p).$$

Taking lim sup of both sides gives

$$\limsup_{n\to\infty} \Lambda_n(x) \le \sup_{p\ge j} x(p).$$

Taking the limit as $j \to \infty$ gives

$$\limsup_{n\to\infty} \Lambda_n(x) \le \limsup_{n\to\infty} x(n).$$

A similar inequality works for limit inferiors, giving us

$$\liminf_{n\to\infty} x(n) \le \Lambda(x) \le \limsup_{n\to\infty} x(n).$$

This is (2).

If $x \in M$, then (1) follows easily (shifting the limit doesn't change anything). Notice that $\tau(x) \in l^\infty$ (shifting still gives a sequence, and it is still bounded). We see that (2) gives us

$$|\Lambda(\tau x) - \Lambda(x)| = |\Lambda(\tau x - x)| \le \limsup_{n\to\infty}(\tau x - x) = 0,$$

since shifting doesn't change the limit superior, so

$$\Lambda(\tau x) = \Lambda(x).$$

□

Problem 24 (Rudin 3.5, James)**.** For $0 < p < \infty$, let l^p be the space of all functions $x : \mathbb{Z}_{\ge 0} \to \mathbb{R}$ (could also be complex) so that

$$\sum_{n=1}^{\infty} |x(n)|^p < \infty.$$

For $1 \le p < \infty$, define

$$\|x\|_p = \left(\sum |x(n)|^p\right)^{1/p},$$
$$\|x\|_\infty = \sup_n |x(n)|.$$

(1) Assume $1 \le p < \infty$. Prove that $\|\cdot\|_p$ and $\|\cdot\|_\infty$ make l^p and l^∞ into Banach spaces.
(2) If $(p,q) = 1$, prove that $(l^p)^* = l^q$ in the following sense: There is a one-to-one correspondence $\Lambda \leftrightarrow y$ between $(l^p)^*$ and l^q given by

$$\Lambda(x) = \sum x(n)y(n), \qquad x \in l^p.$$

(3) Assume $1 < p < \infty$. Prove that l^p contains sequences that converge weakly but not strongly.
(4) On the other hand, prove that every weakly convergent sequence in l^1 converges strongly.
(5) If $0 < p < 1$, prove that l^p metrized by

$$d(x,y) = \sum_{n=1}^{\infty} |x(n) - y(n)|^p$$

is a locally bounded F-space which is not locally convex but that $(l^p)^*$ nevertheless separates points on l^p.
(6) Show that $(l^p)^* = l^\infty$ in the same sense as prior for $0 < p < 1$.
(7) Show also that the set of all x with $\sum |x(n)| < 1$ is weakly bounded but not originally bounded in l^p for $0 < p < 1$.
(8) For $0 < p \le 1$, let τ_p be the weak* topology induced on l^∞ by l^p. If $0 < p < r \le 1$, show that τ_p and τ_r are different topologies but that they induce the same topology on each norm bounded subset of l^∞.

Proof.

(1) We first check that these are indeed norms. The fact that they are positive is clear. We then check the remaining three properties for $\|\cdot\|_p$.

(a) To see the triangle inequality requires a few steps. First, recall that if $a, b \geq 0$ and $0 < \lambda < 1$, then

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b.$$

The proof of this fact is a simple calculus trick (see **Folland Lemma 6.1**). Next, we wish to show that

$$\|xy\|_1 \leq \|x\|_p\|y\|_q$$

for $(p, q) = 1$. Assume that $x, y \neq 0$, since otherwise this is an easy observation. Assume as well $\|x\|_p < \infty$, $\|y\|_q < \infty$, since otherwise this isn't interesting. Normalizing with respect to $\|x\|_p$, $\|y\|_q$, we can assume that these are both 1. Write $a = |x(n)|^p$, $b = |y(n)|^q$, $\lambda = p^{-1}$ to get

$$|x(n)y(n)| = a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b = p^{-1}|x(n)|^p + q^{-1}|y(n)|^q.$$

Adding both sides from $n = 1$ to ∞ gives

$$\|xy\|_1 \leq p^{-1}\|x\|_p^p + q^{-1}\|y\|_q^q = 1 = \|x\|_p\|y\|_q.$$

This gives us Hölder's inequality for sequences.

The goal now is to establish the triangle inequality. Notice that we can assume $x+y \neq 0$. Assume as well $p > 1$. Write

$$|x+y|^p \leq (|x| + |y|)|x+y|^{p-1}$$

and apply Hölder to get

$$\|x+y\|_p^p \leq \|x\|_p\||x+y|^{p-1}\|_q + \|y\|_p\||x+y|^{p-1}\|_q = (\|x\|_p + \|y\|_p)\left(\sum |x(n)+y(n)|^p\right)^{1/q}.$$

Dividing then gives the desired result.

If $p = 1$, then

$$\|x+y\|_1 = \sum |x(n)+y(n)| \leq \sum[|x(n)| + |y(n)|] = \|x\|_1 + \|y\|_1.$$

So we have the triangle inequality for $1 \leq p < \infty$. For $p = \infty$, we see

$$\|x+y\|_\infty = \sup_{n\geq 1}|x(n)+y(n)| \leq \sup_{n\geq 1}|x(n)| + \sup_{n\geq 1}|y(n)| = \|x\|_\infty + \|y\|_\infty.$$

This gives us the triangle inequality for $1 \leq p \leq \infty$.

(b) Homogeneity is an easy observation. For $1 \leq p < \infty$, we have

$$\|\lambda x\|_p = \left(\sum |\lambda x(n)|^p\right)^{1/p} = |\lambda|\left(\sum |x(n)|^p\right)^{1/p} = |\lambda|\|x\|_p.$$

For $p = \infty$ it is equally as easy,

$$\|\lambda x\|_\infty = \sup_{n\geq 1}|\lambda x(n)| = |\lambda|\sup_{n\geq 1}|x(n)|.$$

(c) For $p = \infty$, we see

$$\|x\|_\infty = 0 \iff |x(n)| = 0 \text{ for all n} \iff x = 0.$$

For $1 \leq p < \infty$, we see

$$\|x\|_p = \left(\sum |x(n)|^p\right)^{1/p} \geq (|x(n)|^p)^{1/p} = |x(n)|,$$

so

$$\|x\|_p = 0 \iff |x(n)| = 0 \text{ for all n} \iff x = 0.$$

So these are indeed norms.

To be a Banach space, we need to show that the norm is complete. Fix $1 \le p < \infty$ first. Let

$$\sum \|x_k\|_p$$

be a convergent series. Then the goal is to show

$$\sum x_k$$

converges as well (using **Folland Theorem 5.1**). Define

$$F_N = \sum_{k=1}^{N} |x_k|, \qquad F = \sum_{k=1}^{\infty} |x_k|.$$

Notice

$$\|F_N\|_p \le \sum_{k=1}^{N} \|x_k\|_p \le \sum \|x_k\|_p < \infty$$

for all N. Hence,

$$\lim_{N\to\infty} \|F_N\|_p = \|F\|_p < \infty.$$

In particular, $\|F\|_p < \infty \implies F(n) < \infty$ for all n, so

$$G(n) = \sum_{k=1}^{\infty} x_k(n)$$

converges for all n. The goal is to show

$$\lim_{n\to\infty} \sum_{k=1}^{n} x_k \to G$$

with respect to the L^p norm. Notice

$$\left\| G - \sum_{k=1}^{n} x_k \right\|_p^p = \left\| \sum_{k=n+1}^{\infty} x_k \right\|_p^p \to 0,$$

so we have convergence.

Note that $x_n \to x$ in the l^∞ norm iff

$$\|x_n - x\|_\infty \to 0.$$

In other words,

$$\lim_{n\to\infty} \sup_{k\ge 1} |x_n(k) - x(k)| = 0.$$

This means that $x_n \to x$ uniformly.

Let (x_n) be Cauchy. Then for all $\epsilon > 0$, there exists an N so that for $n, m \ge N$,

$$\|x_n - x_m\|_\infty < \epsilon.$$

This means that for k fixed, we have

$$|x_n(k) - x_m(k)| < \epsilon,$$

so that $\{x_n(k)\}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, we have that a limit is defined. Thus, we can let

$$x(k) = \lim_{n\to\infty} x_n(k)$$

For all $\epsilon > 0$, we wish to show that there is an N so that for $n \ge N$,

$$\|x_n - x\|_\infty < \epsilon.$$

This follows from just noting that that the sequence is Cauchy, so there is an N with $n, m \geq N$ implying

$$\|x_n - x_m\|_\infty < \epsilon,$$

and then we take the limit as $m \to \infty$. This makes l^p into a Banach space for $1 \leq p \leq \infty$.

(2) Recall $(p, q) = 1$ if and only if $1/p + 1/q = 1$. One direction is easy. We see that $y \in l^q$ gives us a linear functional which is continuous by Hölder, since

$$\|xy\|_1 \leq \|x\|_p \|y\|_q < \infty.$$

So $\Lambda : l^p \to \mathbb{R}$ defined by

$$\Lambda(x) = \sum x(n)y(n) < \infty,$$

and continuity follows by noting

$$\|\Lambda\| \leq \|y\|_q.$$

We now consider $\Lambda : l^p \to \mathbb{R}$ a continuous linear functional. The goal is to define a function $y \in l^q$ so that

$$\Lambda(x) = \sum x(n)y(n).$$

Consider

$$x_k(n) = \begin{cases} 1 \text{ if } k = n \\ 0 \text{ otherwise.} \end{cases}$$

We see that $x_k(n) \in l^p$ clearly, and we note that $x \in l^p$ is of the form

$$x(n) = \sum_{k=1}^{\infty} x(k)x_k(n),$$

Now define

$$y(k) = \Lambda(x_k).$$

The claim is that

$$\Lambda(x) = \sum x(n)y(n)$$

and $y \in l^q$. To see the first result, define

$$S_N = \sum_{k=1}^{N} a_k x_k,$$

so that $\lim_{N\to\infty} S_N = x$. Notice

$$\Lambda(S_N) = \sum_{k=1}^{N} x(k)\Lambda(x_k),$$

and by continuity we get

$$\sum_{k=1}^{\infty} x(k)y(k) = \lim_{N\to\infty} \Lambda(S_N) = \Lambda(x).$$

So the first property is established. Next, $\|\Lambda\| < \infty$, since Λ is continuous. Recall

$$\|\Lambda\| = \sup\{|\Lambda(x)| : \|x\|_p = 1\}.$$

The goal is to show $\|y\|_q \leq \|\Lambda\|$. Let $g_n = \chi_{[1,n]}y$, so that $g_n \to y$ pointwise. Let

$$y_n = \frac{|g_n|^{q-1}\overline{\text{sgn}(g)}}{\|g_n\|_q^{q-1}},$$

then $\|y_n\|_p = 1$ and using Fatou's Lemma we get

$$\|y\|_q \leq \liminf \|y_n\|_q \leq \liminf \sum y_n(k)y(k) \leq \|\Lambda\|.$$

This gives us the desired correspondence.

(3) Assume $1 < p < \infty$. Consider $x_k(n)$ given above. We wish to show that $x_k(n) \to 0$ weakly but not strongly. Notice that for any k and m,

$$\|x_k - x_m\|_p = 2^{1/p}.$$

The sequence is not Cauchy, so does not converge strongly to anything. On the other hand, for any $y \in l^q$, we have

$$\sum_{n=1}^{\infty} x_k(n)y(n) = y(k).$$

Since $\|y\|_q < \infty$, we must have $\lim_{k\to\infty} |y(k)| = 0$. Hence,

$$\lim_{k\to\infty} \Lambda(x_k) = 0 \text{ for all } \Lambda \in (l^p)^*.$$

So it converges weakly to 0.

(4) Suppose $x_n \to x$ weakly in l^1. Without loss of generality, we assume that $x_n \to 0$ weakly by linearity. For contradiction, assume $x_n \not\to 0$. We follow the method of the "gliding hump" (credit to this user) Since $x_n \not\to 0$, there is a subsequence (x_m) so that $\|x_m\|_1 \geq \epsilon$. The idea is to construct a linear functional so that

$$f(x_n) = \sum x_n(k)y(k) > 0 \text{ for all n.}$$

Choose N_1 so that

$$\sum_{k=N_1}^{\infty} |x_1(k)| < \frac{\epsilon}{100}.$$

Choose M_1 so that

$$\sum_{k=1}^{N_1} |x_{M_1}(k)| < \frac{\epsilon}{100}.$$

Choose corresponding N_j and M_j so that

$$\sum_{k=N_j}^{\infty} |x_{M_{j-1}}(k)| < \frac{\epsilon}{100},$$

$$\sum_{k=1}^{N_j} |x_{M_j}(k)| < \frac{\epsilon}{100}.$$

Given M_{j-1} (and setting $M_0 = 1$) we see we can choose N_j since each $x_{M_{j-1}} \in l^1$. Given N_j, we see we can choose M_j since the sequence $y(k) = 1$ for $1 \leq k \leq N_j$ is in l^∞ which is contained in the dual, so there must be some M_j which satisfies the criteria by weak convergence.

Notice that the bulk of the mass for x_{M_j} lies in between $N_{j-1} + 1 \leq k \leq N_j$, since

$$\sum_{k=1}^{N_{j-1}} |x_{M_{j-1}}(k)| < \frac{\epsilon}{100},$$

$$\sum_{k=N_j}^{\infty} |x_{M_{j-1}}(k)| < \frac{\epsilon}{100},$$

and $\|x\|_1 > \epsilon$. Define

$$y(k) = \text{sgn}(x_{M_j}(k)) \text{ for } N_{j-1} < k \leq N_j.$$

Then we see that the linear functional defined by y gives

$$|f(x_{M_j})| = \left|\sum_{k=1}^{\infty} y(k)x_{M_j}(k)\right| = \left|\sum_{k=1}^{N_{j-1}} y(k)x_{M_j}(k) + \sum_{k=N_{j-1}+1}^{N_j} y(k)x_{M_j}(k) + \sum_{k=N_j+1}^{\infty} y(k)x_{M_j}(k)\right|$$
$$\geq \frac{98}{100}\epsilon > 0.$$

This applies for an infinite subsequence x_{M_j}, so this contradicts the fact that $x_n \to 0$ weakly.

(5) Recall a space is said to be *locally bounded* if the origin has a set which is bounded. A set $E \subset X$ is said to be *bounded* if for all neighborhoods of the origin V there exists a $t > 0$ with $E \subset tV$. A space is an *F-space* if the topology is induced by a complete invariant metric.

We need to show the following:

(a) The function d is indeed a metric.
(b) The metric d is invariant.
(c) The metric d is complete (this shows that it is an F-space).
(d) The origin admits a bounded neighborhood (this shows that it is a locally bounded F-space).
(e) The space is not locally convex.
(f) $(l^p)^*$ separates points on l^p.

We go through each step:

(a) We show that d is a metric. There are four properties to check.
  (i) It is clear that $d(x, y) \geq 0$, since it is a sum of positive terms.
  (ii) Notice that
$$d(x, y) = \sum |x(n) - y(n)|^p = \sum |y(n) - x(n)|^p = d(y, x).$$
  (iii) We see $d(x, y) = 0$ iff $|x(n) - y(n)|^p = 0$ for each n iff $x(n) = y(n)$ for each n iff $x = y$.
  (iv) The triangle inequality is easier this time. We claim for $a, b \geq 0$, we have
$$(a + b)^p \leq a^p + b^p$$
for $0 < p < 1$. If $b = 0$, it is clear. Assume $b \neq 0$. Divide both sides by b^p to get
$$\left(\frac{a}{b} + 1\right)^p \leq \left(\frac{a}{b}\right)^p + 1.$$
Let $t = a/b$ (since $b \neq 0$). Then the goal is to show for $t \geq 0$,
$$(t + 1)^p \leq t^p + 1.$$
In other words, the goal is to show
$$(t + 1)^p - t^p \leq 1.$$
Taking the derivative of the left hand side with respect to t, we get
$$p(t + 1)^{p-1} - pt^{p-1}.$$
For $t > 0$ we get that this is less than 0 so that the function is decreasing. at $t = 0$, we get 1, so it is maximized at 1. In other words, we have the desired inequality.

With this in mind, we see that for each n

$$|x(n)-y(n)|^p = |x(n)-z(n)+z(n)-y(n)|^p \leq (|x(n)-z(n)|^p + |z(n)-y(n)|^p,$$

so

$$d(x,y) \leq d(x,z) + d(y,z).$$

This establishes that d is a metric.

(b) To see invariance, notice

$$d(x+z, y+z) = \sum |x(n)+z(n)-y(n)-z(n)|^p = \sum |x(n)-y(n)|^p = d(x,z).$$

(c) To see that it's complete, let (x_n) be a Cauchy sequence with respect to the metric. In other words, for all $\epsilon > 0$ there exists an N so that for $n, m \geq N$ we have

$$d(x_n, x_m) < \epsilon.$$

Translating,

$$\sum |x_n(k) - x_m(k)|^p < \epsilon.$$

For fixed k, we have

$$|x_n(k)-x_m(k)|^p < \epsilon \implies |x_n(k)-x_m(k)| < \epsilon^{1/p}.$$

Since $0 < p < 1$ is fixed and this holds for all $\epsilon > 0$, we get that the sequence $(x_n(k))$ is Cauchy in $\mathbb{R}$. Utilizing the fact that $\mathbb{R}$ is complete, we get that $(x_n(k))$ converges to a value, denote it by $x(k)$.

The idea now is that $x_n \to x$. To see this, fix $\epsilon > 0$. Since the sequence is Cauchy, there exists an N so that for $n, m \geq N$, we have

$$d(x_n, x_m) < \frac{\epsilon}{2}.$$

Notice that for $n \geq N$, we have

$$d(x, x_n) \leq d(x, x_m) + d(x_n, x_m).$$

Notice

$$d(x, x_m) = \sum |x(k) - x_m(k)|^p.$$

For sufficiently large m, we get

$$|x(k) - x_m(k)|^p < \frac{\epsilon}{2^{k+1}},$$

so

$$d(x, x_m) < \frac{\epsilon}{2}.$$

This gives us $d(x, x_n) < \epsilon$. The choice of $\epsilon > 0$ was arbitrary, so $x_n \to x$. This establishes completeness.

(d) The balls in this topology are

$$B_n = \{x \in l^p : d(x, 0) < 1/n\}.$$

Notice that

$$n^{1/p} B_n = \{n^{1/p} x \in l^p : d(x,0) < 1/n\}.$$

Notice

$$d(n^{1/p}x, 0) = \sum |n^{1/p} x(k)|^p = n \sum |x(k)|^p < 1.$$

So $B_1 = n^{1/p} B_n$, establishing B_1 is bounded.

(e) Fix $\epsilon > 0$. If l^p was locally convex, then we would have $\text{co}(B_\epsilon) \subset B_\delta$ for some $\delta > 0$ (where here co is the convex hull). Notice that for $k \geq 1$, $m \geq 1$ we can set

$$x_k(n) = \begin{cases} \frac{\epsilon}{m} \text{ if } n = k \\ 0 \text{ otherwise.} \end{cases}$$

Then

$$(x_1 + \cdots + x_m) \in \text{co}(B_\epsilon).$$

However,

$$\|x_1 + \cdots + x_m\|_p = \frac{\epsilon^p}{m^p} + \cdots + \frac{\epsilon^p}{m^p} = m^{1-p}\epsilon^p,$$

so if this were in B_δ for some $\delta > 0$, we require

$$m^{1-p}\epsilon^p < \delta \implies m^{1-p} < \frac{\delta}{\epsilon^p}$$

for all $m \geq 1$. Since $p < 1$, $1 - p > 0$, so it is impossible for this to hold for all m. Thus, it is impossible for $\text{co}(B_\epsilon) \subset B_\delta$ for any δ.

(f) If $x \neq y$, then there is a k so that $x(k) \neq y(k)$. Let $\Lambda : l^p \to \mathbb{R}$ be defined by $\Lambda(x) = x(k)$. This is a linear functional, since $\Lambda(x+y) = (x+y)(k) = x(k)+y(k) = \Lambda(x)+z(y)$, and $\Lambda(ax) = (ax)(k) = ax(k)$. It is continuous, since $\mathcal{N}(\Lambda) = \{x : x(k) = 0\}$ is closed (here utilizing **Theorem 1.18**). If $(x_n) \subset N(\Lambda)$, $x_n \to x$, then we have that

$$x_n(k) = 0 \text{ for all } k,$$

so for all $\epsilon > 0$ we get

$$|x(k)|^p < \epsilon.$$

Hence $x(k) = 0$, so $x \in N(\Lambda)$. So Λ is a continuous linear functional with $\Lambda(x) \neq \Lambda(y)$. We can find one for each $x \neq y$, so $(l^p)^*$ separates points.

(6) Let $y \in l^\infty$. We wish to show

$$\sum x(k)y(k) < \infty,$$

so that y defines a linear functional by

$$\Lambda(x) = \sum x(k)y(k).$$

Let

$$S_N = \sum_{k=0}^{N} x(k)y(k).$$

Inducting the lemma from prior, we have

$$|S_N|^p \leq \sum_{k=0}^{N} |x(k)|^p |y(k)|^p.$$

Since $y \in l^\infty$, $|y(k)| \leq M$ for some $M < \infty$. So

$$|S_N|^p \leq M^p \sum_{k=0}^{N} |x(k)|^p.$$

Taking the limit as $N \to \infty$, we get

$$|S|^p \leq M^p \sum_{k=0}^{\infty} |x(k)|^p,$$

where

$$S = \sum x(k)y(k).$$

This establishes that $S < \infty$, so that Λ is well-defined. Linearity is an easy observation, and continuity follows by the fact that Λ is bounded on B_1, the ball of radius 1 around the origin. So $l^\infty \subset (l^p)^*$.

For the other direction, let $\Lambda : l^p \to \mathbb{R}$ be a continuous linear functional. Again, define

$$x_k(n) = \begin{cases} 1 \text{ if } n = k \\ 0 \text{ otherwise.} \end{cases}$$

For $x \in l^p$, we can write

$$x(n) = \sum a_k x_k(n).$$

By the same argument as before, we have

$$\Lambda(x) = \sum a_k \Lambda(x_k).$$

Let $y(k) = \Lambda(x_k)$. We claim $y \in l^\infty$. Since Λ is continuous, we have that there is some ball B_ϵ around the origin with Λ bounded on B_ϵ. So for sequences $x \in B_\epsilon$, we get $|\Lambda(x)| \leq M$. Consider the sequence

$$x_k(n) = \begin{cases} \epsilon^{1/p} \text{ if } k = n \\ 0 \text{ otherwise.} \end{cases}$$

Then $x_k(n) \in B_\epsilon$, and $|\Lambda(x_k)| = \epsilon^{1/p}|y(k)| \leq M$. So for all k, we have

$$|y(k)| \leq \frac{M}{\epsilon^{1/p}} \implies \|y\|_\infty < \infty.$$

So $y \in l^\infty$, as desired. We then get the correspondence.

(7) Consider

$$E = \left\{ x \in l^p : \sum |x(n)| < 1 \right\}.$$

We wish to show that for all $y \in l^\infty$, $x \in E$, we have

$$\sum |x(k)||y(k)| < \infty.$$

But this is clear, since

$$|y(k)| \leq M \text{ for all } k,$$

so

$$\sum |x(k)||y(k)| \leq M \sum |x(k)| < \infty.$$

Hence, E is weakly bounded, since every $\Lambda \in (l^p)^*$ is bounded on E.

We now want to show it is not originally bounded. So for all $t > 0$, it suffices to show that $E \not\subset tB_1$, where B_1 is the ball of radius 1. Note

$$tB_1 = \left\{ tx \in l^p : \sum |x(n)|^p < 1 \right\} = \left\{ x \in l^p : \sum |x(n)|^p < t^p \right\}.$$

So for each $t > 0$, it suffices to find $x \in l^p$ so that

$$\sum |x(n)| < 1$$

but

$$\sum |x(n)|^p \geq t^p.$$

For $0 < t < 1$, this is easy. Set $x(1) = t$ and $x(k) = 0$ otherwise. Since $0 < p < 1$, we get $1/p > 1$. So the series

$$\sum n^{-1/p} = \gamma(p) < \infty.$$

Let $x(n) = \delta n^{-1/p}$ for $\delta < 1/\gamma(p)$ and $n \leq N$, $x(n) = 0$ for $n > N$. Here, N is chosen sufficiently large so that

$$\sum_{n=0}^{N} |x(n)|^p \geq t^p.$$

Note that such an N exists since the series diverges. Then we see that

$$\sum |x(n)| < 1,$$

while

$$\sum |x(n)|^p \geq t^p$$

by construction. So we've found $x \in E$ with $x \notin tB_1$ for all $t > 0$, so the set is not bounded.

(8) We now consider the weak* topology generated on l^∞ by l^p for $0 < p \leq 1$. We can write $x \in l^p$ as a linear functional on l^∞ by

$$x(y) = \sum x(k)y(k).$$

For a finite collection $\{x_i\}_{i=1}^n \subset l^p$, consider

$$W = \{y \in l^\infty : |x_i(y)| < \epsilon_i, 1 \leq i \leq n\}$$

Sets of the form W form the open balls for the weak* topology on l^∞. For $0 < r < p \leq 1$, we can take $\{z_j\}_{j=1}^m \subset l^r$ and write

$$V = \{y \in l^\infty : |z_j(y)| < \epsilon_i, 1 \leq j \leq m\}.$$

Note that these form the open balls in the weak* topology induced by l^r. The goal is to show we can form an open ball with respect to one that is not open with respect to the other.

Take $z \in l^r$. Then we have

$$\sum |z(k)|^r = M < \infty.$$

Notice

$$\left(\sum |z(k)|^p\right)^{r/p} \leq \sum |z(k)|^{p(r/p)} = \sum |z(k)|^r = M.$$

So

$$\sum |z(k)|^p = M^{p/r} < \infty.$$

Hence $l^r \subset l^p$. This inclusion is strict, so we see that the topologies generated differ (the topology generated by τ_r will be weaker, since there are less functions).

It suffices to show it on the norm-closed unit ball. Banach-Alaoglu says this is weak* compact with respect to both τ_r and τ_p. These are both Hausdorff topologies, so since all compact Hausdorff topologies must agree, we have that $\tau_p = \tau_r$ on the unit ball.

□

Remark. Reference for last fact found here.

Problem 25 (Rudin 3.11, James)**.** Let X be an infinite-dimensional Fréchet space. Prove that X^* with its weak* topology is of the first category in itself.

Proof. Recall that X^* is of the first category if it is a countable union of nowhere dense sets. We can embed X into X^{**} via $J : X \to X^{**}$ with $J(x)(\varphi) = \varphi(x)$; i.e., $J(x) = \text{Eval}_x$. Then the weak* topology says that any continuous linear functional on X^* is of the form $J(x)$ for some $x \in X$.

Recall that $\{x_1, \ldots, x_n\} \subset X$, $\epsilon_i > 0$, sets of the form

$$V = \{\varphi \in X^* : |J(x_i)(\varphi)| < \epsilon_i \text{ for } 1 \leq i \leq n\} = \{\varphi \in X^* : |\varphi(x_i)| < \epsilon_i \text{ for } 1 \leq i \leq n\}$$

give us a local base for our topology.

Consider U a neighborhood of 0 in a topological vector space X. We define the *polar* of U to be

$$\mathrm{Pol}(U) = \{\varphi \in X^* : |\varphi(x)| \le 1 \text{ for all } x \in V\}.$$

We see that $\mathrm{Pol} : \mathcal{P}(X) \to \mathcal{P}(X^*)$ satisfies a reverse inclusion property. To see this, let $U \subset V$ be neighborhoods of the origin, and take $\varphi \in \mathrm{Pol}(V)$. Then

$$|\varphi(x)| \le 1 \text{ for all } x \in V \implies |\varphi(x)| \le 1 \text{ for all } x \in U,$$

so $\varphi \in \mathrm{Pol}(U)$. Hence, $\mathrm{Pol}(V) \subset \mathrm{Pol}(U)$.

Since X is a Frèchet space, we can take a sequence of decreasing open balls $\{U_n\}$. This corresponds to a sequence of increasing polar sets $\{\mathrm{Pol}(U_n)\}$ with

$$X^* \subset \bigcup_n \mathrm{Pol}(U_n).$$

The proof of Banach-Alaoglu says that $\mathrm{Pol}(U_n)$ is a weak*-closed set for each n. So if we can show $\mathrm{Pol}(U_n)^{\mathrm{o}} = \varnothing$, we win. If it was nonempty, we have some V given above so that $V \subset \mathrm{Pol}(U_n)$, since these are basic open sets. In other words, for some small ϵ, we get that

$$\{\varphi \in X^* : |\varphi(x_i)| < \epsilon \text{ for } 1 \le i \le n\} \subset \{\varphi \in X^* : |\varphi(x)| \le 1 \text{ for all } x \in U_n\}.$$

For each $x \in U_n$, we get $\ker(J(x)) \subset \bigcap_{i=1}^n \ker(J(x_i))$, so by **Rudin Lemma 3.9** we get x is a linear combination of the x_i. So U_n is a finite dimensional vector subspace. The U_n are bounded though, so this implies that $U_n = \{0\}$, which is a contradiction. So it is in fact impossible to find V so that $V \subset \mathrm{Pol}(U_n)$, implying that the interior of $\mathrm{Pol}(U_n)$ is trivial. □

Remark. The gist came from here.

Problem 26 (Rudin 3.18, James)**.** Let K be the smallest convex set in $\mathbb{R}^3$ containing the points $(1, 0, 1)$ $(1, 0, -1)$, and $(\cos(\theta), \sin(\theta), 0)$ for $0 \le \theta < 2\pi$.

(1) Show that K is compact.
(2) Show that the set of all extreme points of K is not compact.
(3) Does such an example exist in $\mathbb{R}^2$.

Proof. (1) Let

$$A := \{(1, 0, 1)\} \cup \{(1, 0, -1)\} \cup \{(\cos(\theta), \sin(\theta), 0) : 0 \le \theta < 2\pi\}.$$

By **Rudin Theorem 3.20**, we have that if $A \subset \mathbb{R}^3$ is compact, then $K = \mathrm{co}(A)$ is compact. Notice

$$A \subset \overline{B_2(0)},$$

so if we show A is closed then A is compact. We see that

$$\begin{aligned} \overline{A} &= \overline{\{(1, 0, 1)\}} \cup \overline{\{(1, 0, -1)\}} \cup \overline{\{(\cos(\theta), \sin(\theta), 0) : 0 \le \theta < 2\pi\}} \\ &= \{(1, 0, 1)\} \cup \{(1, 0, -1)\} \cup \overline{\{(\cos(\theta), \sin(\theta), 0) : 0 \le \theta < 2\pi\}}. \end{aligned}$$

So it suffices to show this last set is closed. Let

$$(x_n) \subset \{(\cos(\theta), \sin(\theta), 0) : 0 \le \theta < 2\pi\}.$$

We can express these as

$$x_n = (\cos(\theta_n), \sin(\theta_n), 0).$$

Suppose $x_n \to x \in \mathbb{R}^3$. This means

$$(\cos(\theta_n), \sin(\theta_n), 0) \to x.$$

Since these are continuous functions, this forces $\theta_n \to \theta \in [0, 2\pi)$, so

$$x = (\cos(\theta), \sin(\theta), 0).$$

Hence the set is closed, so A is closed and compact.

(2) Denote $E(K)$ to be the set of extreme points. These are certainly bounded, so we need to show that the set is not closed. The point $(1,0,0)$ is not in $E(K)$ (since it is the midpoint of the line connecting $(1,0,1)$ and $(1,0,-1)$) but $(\cos(\theta),\sin(\theta),0)$ for $0<\theta<2\pi$ are extreme points. The set is not closed then.

(3) It does not. A simple (geometric) proof by contradiction shows that extreme points are isolated here.

□

Problem 27 (Rudin 3.24, James)**.** Suppose

(1) X is a topological vector space on which X^* separates points.
(2) Y is a topological vector space on which Y^* separates points.
(3) μ is a Borel probability measure on a compact Hausdorff space Q.
(4) $f: Q \to X$ is continuous.
(5) $\overline{\text{co}}(f(Q))$ is compact.
(6) $T: X \to Y$ a continuous linear mapping.

Prove

$$T\left(\int_Q f d\mu\right) = \int_Q (Tf) d\mu.$$

Proof. Take $\Lambda \in Y^*$. Then we claim that $\Lambda \circ T \in X^*$. This follows since composition of linear functions is linear, composition of continuous functions is continuous. By **Theorem 3.27**, we see that there is a $y \in X$ with

$$y = \int_Q f d\mu.$$

Notice this y satisfies the property that for every $\Lambda \in X^*$, we have

$$\Lambda(y) = \int_Q \Lambda(f) d\mu.$$

Now we examine

$$T(y) = T\left(\int_Q f d\mu\right).$$

Taking any linear function $\Lambda \in Y^*$, we have

$$\Lambda \circ T(y) = \Lambda \circ T\left(\int_Q f d\mu\right) = \int_Q \Lambda \circ T(f) d\mu.$$

Since this holds for all $\Lambda \in Y^*$, we get by definition that

$$T(y) = \int_Q T(f) d\mu.$$

□

1.4. **Chapter 4.**

Problem 28 (Rudin 4.1, James)**.** Let φ be the embedding of X into X^{**}. Let τ be the weak topology of X, and let σ be the weak* topology of X^{**}.

(1) Prove that φ is a homeomorphism of (X,τ) onto a dense subspace of X^{**}.
(2) If B is the closed unit ball of X, prove that $\varphi(B)$ is σ-dense in the closed unit ball of X^{**}.
(3) Use (1), (2), and Banach-Alaoglu to prove that X is reflexive iff B is weakly compact.
(4) Deduce from (3) that every norm-closed subspace of a reflexive space is reflexive.
(5) If X is reflexive and Y is a closed subspace of X, prove that X/Y is reflexive.

(6) Prove that X is reflexive iff X^* is reflexive.

Proof.

(1) We have that $\varphi(x) : X^* \to \mathbb{F}$ is defined by $\varphi(x)(f) = f(x)$. We break this up into parts.
  (a) The mapping φ is injective: This follows by noting that $\varphi(x) = \varphi(y)$ if and only if for all $x^* \in X^*$, $x^*(x) = x^*(y)$. Since X^* separates points, this forces $x = y$.
  (b) Consider a weak* open neighborhood in X^{**}, i.e. consider
$$V = \{x^{**} \in X^{**} : |x^{**}(x_i^*)| < \epsilon \text{ for } 1 \le i \le n\}$$
  where $(x_i^*)_{i=1}^n \subset X^*$. Taking the preimagine of V with respect to φ tells us
$$\begin{aligned}\varphi^{-1}(V) &= \{x \in X : |\varphi(x)(x_i^*)| < \epsilon \text{ for } 1 \le i \le n\} \\ &= \{x \in X : |x_i^*(x)| < \epsilon \text{ for } 1 \le i \le n\}.\end{aligned}$$
  This is open with respect to the weak topology on X. This holds for all open basic sets, so φ continuous.

  Hence φ is a homeomorphism onto its image. The fact $\operatorname{Im}(\varphi)$ is dense follows from Goldstine's theorem.

(2) This is Goldstine's theorem.

(3) ($\implies$): Assume X is reflexive, so that $X^{**} = \varphi(X)$. We have $B = B^{**}$, and by Banach-Alaoglu B^{**} is weak* compact. So by (1) we get that B is weakly compact.
  ($\impliedby$): If B is weakly compact, $\varphi(B) \subset B^{**}$ is weakly dense, Since B is weakly closed, this implies $\varphi(B) = B = B^{**}$. By scaling, we get $X = X^{**}$.

(4) Let $Y \subset X$ be a norm-closed subspace, X reflexive. By (3), B is weakly compact, hence $B \cap Y$ is weakly compact, and this is the closed unit ball in Y. So Y is reflexive.

(5) X/Y is a norm-closed subspace of X.

(6) X is reflexive implies X^* is reflexive is easy. For the other direction, assume X^* is reflexive. X is a norm-closed subspace of X^{**}, so in particular it is closed in the topology induced on X^{**} by X^{***}. But $X^{***} = X^*$, so it is weak* closed. The image $\varphi(X) = X \subset X^{**}$ is weak* dense, so in particular $X = X^{**}$. So X is reflexive.

□

Remark. This was helpful in the last theorem. In particular, **Theorem 6.28, 6.29**.

Problem 29 (Rudin 4.2, James)**.**

(1) Which of the following spaces are reflexive?
  (a) c_0.
  (b) ℓ^1.
  (c) ℓ^p, $1 < p < \infty$.
  (d) ℓ^p, $0 < p < 1$.
  (e) ℓ^∞.

(2) Prove that every finite dimensional normed space is reflexive.

(3) Prove that C, the supremum-normed space of all complex continuous functions on the unit interval is not reflexive.

Proof.

(1) (a) Recall
$$c_0 = \{x : \mathbb{N} \to \mathbb{R} : \lim_{n\to\infty} |x(n)| = 0\}.$$
  I can't remember if we established the dual of c_0, so let's do that now. Let
$$z_k(n) = \begin{cases} 1 \text{ if } n = k \\ 0 \text{ otherwise.} \end{cases}$$

For $x \in c_0$, we can express it as

$$x(n) = \sum_{k=1}^{\infty} a_k z_k(n),$$

Now, by prior exercises, for $y^* \in c_0^*$ we see that

$$y^*(x) = \sum_{k=1}^{\infty} a_k y^*(z_k) = \sum_{k=1}^{\infty} x(k) y^*(z_k).$$

Define $y : \mathbb{N} \to \mathbb{R}$ by $y(k) = y^*(z_k)$. Then we can write the above as

$$y^*(x) = \sum_{k=1}^{\infty} x(k) y(k).$$

The claim now is that $y \in \ell^1$. Since the above holds for all $x \in c_0$, define

$$x(k) = \begin{cases} \frac{|y(k)|}{y(k)} \text{ if } y(k) \neq 0 \\ 0 \text{ otherwise.} \end{cases}$$

Then

$$y^*(x) = \sum_{k=1}^{\infty} |y(k)| < \infty.$$

We've seen from before that every $y \in \ell^1$ can be identified with a linear functional on c_0, so we have that $\ell^1 = c_0^*$. Now recall that $(\ell^1)^* = \ell^\infty \neq c_0$. So c_0 is not reflexive.

(b) We have $(\ell^1)^* = \ell^\infty$, $(\ell^\infty)^* \neq \ell^1$, so ℓ^1 is not reflexive. To see this last fact of $(\ell^\infty)^* \neq \ell^1$, define $f(x) = \lim_n x_n$. This is a linear functional on $(\ell^\infty)^*$, but cannot be identified with an element in ℓ^1. See here.

(c) We saw before that $(\ell^p)^* = \ell^q$, where $(p, q) = 1$, and hence $(\ell^q)^* = \ell^p$. So this is reflexive.

(d) We saw before that $(\ell^p)^* = \ell^\infty$, $(\ell^\infty)^* \neq \ell^p$.

(e) Recall X is reflexive iff X^* is reflexive, so ℓ^1 not being reflexive forces ℓ^∞ to not be reflexive.

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

(3) Consider

$$C = \{f : [0,1] \to \mathbb{C} : f \text{ is continuous.}\}$$

Apply the extreme points argument here to get that C is not the dual of a vector space, hence can't be reflexive.

□

Problem 30 (Rudin 4.13, James)**.**

(1) Suppose $T \in B(X, Y)$, $T_n \in B(X, Y)$ for $n \geq 1$, each T_n has finite dimensional range, and

$$\lim \|T - T_n\| = 0.$$

Prove that T is compact.

(2) Assume Y is a Hilbert space, and prove the converse of (1): Every compact $T \in B(X, Y)$ can be approximate in the operator norm by operators with finite-dimensional ranges.

Proof.

(1) Let
$$U = \{x \in X : \|x\| < 1\}$$
be the open ball in X. Recall $T : X \to Y$ is said to be *compact* if $\overline{T(U)} \subset Y$ is compact in Y. Equivalently, T is compact if and only if every bounded sequence $(x_n) \subset X$ contains a subsequence (x_{n_j}) such that (Tx_{n_j}) converges in Y. Equivalently, T is compact if and only if it is totally bounded.

Recall from **Theorem 4.18** that if $T \in B(X,Y)$ and $\dim(\mathrm{Im}(T)) < \infty$ then T is compact. Each T_n has finite-dimensional range, so $\dim(\mathrm{Im}(T_n)) < \infty$, implying that each T_n is compact. The goal now is to show this forces T to be compact.

Let $(x_n) \subset X$ be a bounded sequence. The goal is to show there is a subsequence so that Tx_{n_j} converges. T_1 is compact, so we can refine x_n to $x_n^{(1)}$ (a subsequence) so that $T_1 x_n^{(1)}$ converges. We can then refine this so that $T_2 x_n^{(2)}$ converges. This gives us a sequence of nested sequences $(x_n^{(j)})$. Let $y_n = x_n^{(n)}$ be the diagonal component. We claim that Ty_n converges. We see that for every n we have $T_m y_n$ converges (since eventually n will be large enough to be the converging sequence). We then need to show that Ty_n is Cauchy. To do so, we see that adding and subtracting $T_m y_n$ gives
$$\|T(y_n - y_k)\| \le \|Ty_n - T_m y_n\| + \|T_m y_n - T_m y_k\| + \|T_m y_k - Ty_k\|.$$
Notice that y_n is a bounded sequence, so
$$\|Ty_n - T_m y_n\|, \|T_m y_k - Ty_k\| \le M\|T - T_m\|.$$
Since eventually for large n, k, m we have that the middle term is small, and this is all independent of m, we can take $m \to \infty$ to make everything go to 0.

Remark. This argument actually establishes that if
$$\lim \|T - T_n\| = 0$$
and (T_n) are compact, then T is compact.

(2) Since Y is a Hilbert space, we can find an orthonormal basis $(e_j) \subset Y$. Let P_n be the projection onto $\mathrm{span}\{e_1, \dots, e_n\}$. Define $Q_n = \mathrm{Id} - P_n$. This forces $\|Q_n Tx\|$ to be a decreasing function with respect to n for all x, so $\|Q_n T\|$ is decreasing. The claim now is that $\|T - Q_n T\| \to 0$. If this is the case, then we're done (since that's the only property we're missing now). Assume for contradiction there is a c so that $\|T - Q_n T\| \ge c$. Choose $(x_n) \subset X$ with $\|x\| = 1$ and $\|Q_n T x_n\| \ge c/2$. By compactness of T, we can find a subsequence x_{n_j} so that $Tx_{n_j} \to y$. Then
$$\|Tx_{n_j} - Q_{n_j} T x_{n_j}\| \le \|Q_{n_j} y\| + \|Q_{n_j}(y - Tx_{n_j})\| \le \|Q_{n_j} y\| + \|y - Tx_{n_j}\|.$$
We see that $\|Q_n y\| \to 0$, and so both terms on the right converge to 0. This is the contradiction.

□

Remark. Reference can be found here.

Problem 31 (Rudin 4.15, James)**.** Suppose μ is a finite (or σ-finite) positive measure on a measure space Ω, $\mu \times \mu$ is the corresponding product measure on $\Omega \times \Omega$, and $K \in L^2(\mu \times \mu)$. Define
$$(Tf)(s) = \int_\Omega K(s,t) f(t) d\mu(t). \qquad [f \in L^2(\mu)]$$

(1) Prove that $T \in B(L^2(\mu))$ and that
$$\|T\|^2 \le \int_\Omega \int_\Omega |K(s,t)|^2 d\mu(s) d\mu(t).$$

(2) Suppose a_i, b_i are members of $L^2(\mu)$, for $1 \le i \le n$ put $K_1(s,t) = \sum a_i(s)b_i(t)$ and define T_1 in terms of K_1 as T was defined in terms of K. Prove that $\dim(\mathrm{Im}(T_1)) \le n$.
(3) Deduce that T is a compact operator on $L^2(\mu)$.
(4) Suppose $\lambda \in \mathbb{C}$, $\lambda \neq 0$. Prove: Either the equation

$$Tf - \lambda f = g$$

has a unique solution $f \in L^2(\mu)$ for every $g \in L^2(\mu)$ or there are infinitely many solutions for some g and none for others.
(5) Describe the adjoint of T.

Proof.

(1) To show that $Tf \in L^2(\mu)$, we need to establish that

$$\int_\Omega \left| \int_\Omega K(s,t)f(t)d\mu(t) \right|^2 d\mu(s) < \infty.$$

First, notice that we can bring the absolute values inside to get an upper bound of

$$\int_\Omega \left(\int_\Omega |K(s,t)||f(t)|d\mu(t) \right)^2 d\mu(s).$$

Next, observe that Cauchy-Schwarz tells us that

$$\left(\int_\Omega |K(s,t)||f(t)|d\mu(t) \right)^2 \le \left(\int_\Omega |K(s,t)|^2 d\mu(t) \right) \left(\int |f(t)|^2 d\mu(t) \right).$$

Since $f \in L^2(\mu)$, the value on the right is $\|f\| < \infty$. So we have an upper bound of

$$\|f\| \int_\Omega \int_\Omega |K(s,t)|^2 d\mu(t) d\mu(s).$$

Fubini-Tonelli applies to let us rearrange the order of integration, telling us that we have an upper bound of

$$\|f\| \int_\Omega \int_\Omega |K(s,t)|^2 d\mu(s) d\mu(t).$$

Since $K \in L^2(\mu \times \mu)$, this is finite. This tells us that $Tf \in L^2(\mu)$. The fact that T is linear is an easy thing to see. Furthermore, we have

$$\|T\| = \sup\{\|Tf\|_2 : \|f\|_2 = \|f\| = 1\}$$

$$= \sup\left\{ \sqrt{\int_\Omega \left| \int_\Omega K(s,t)f(t)d\mu(t) \right| d\mu(s)} : \|f\| = 1 \right\}.$$

Squaring both sides, we see that from the above discussion we have an upper bound of

$$\|T\| \le \int_\Omega \int_\Omega |K(s,t)|^2 d\mu(s) d\mu(t).$$

(2) We have

$$(T_1 f)(s) = \int_\Omega K_1(s,t)f(t)d\mu(t).$$

Suppose the (a_i) and (b_i) are basis vectors in $L^2(\mu)$. The fact that $K_1(s,t) \in L^2(\mu \times \mu)$ follows from the fact that

$$|K_1(s,t)|^2 \le \left(\sum |a_i(t)||b_i(s)| \right)^2 \le \sum |a_i(t)|^2 \sum |b_i(s)|^2.$$

This is the finite Cauchy-Schwarz inequality (see here). Integrating this and using Fubini-Tonelli, we have

$$\int_{\Omega\times\Omega} |K_1(s,t)|^2 d\mu(s\times t) = \left(\sum \|a_i\|_2\right)\left(\sum \|b_i\|_2\right) < \infty.$$

We then implement (1) to deduce that $T_1 \in B(L^2(\mu))$. Next, we wish to show that $\dim(\mathrm{Im}(T_1)) < \infty$. We see that

$$\begin{aligned}(T_1 f)(s) &= \int_\Omega K_1(s,t) f(t) d\mu(t) \\ &= \int_\Omega \left(\sum_{i=1}^n a_i(t) b_i(s)\right) f(t) d\mu(t) \\ &= \int_\Omega \sum_{i=1}^n a_i(t) b_i(s) f(t) d\mu(t) \\ &= \sum_{i=1}^n b_i(s) \int_\Omega a_i(t) f(t) d\mu(t) \\ &= \sum_{i=1}^n b_i(s) \langle a_i, f\rangle.\end{aligned}$$

If we take the a_i to all be basis vectors, we see that $\langle a_i, f\rangle = 0$ unless f shares a component with a_i. That is, if f is another basis element, it is 0 unless f is among the a_i. Thus, we have that $\dim(\mathrm{Im}(T_1)) \leq n$.

(3) Use the prior exercise to note that we have a limit of compact operators.

(4) We can rewrite the above to be

$$(T - \lambda e) f = g.$$

If $\lambda \in \sigma(T)$ (so that T is not invertible) then we claim there will be some g so that there are no solutions and some for which there are infinitely many solutions. If $\lambda \notin \sigma(T)$, then the uniqueness and existence of solutions is clear by invertibility.

Note that every $\lambda \in \sigma(T)$ is an eigenvalue by **Theorem 4.25 (b)**. Hence $(T - \lambda e)$ fails to be injective, so 0 has infinitely many solutions. We can then invoke **Theorem 4.24** to see that this also fails to be surjective. So some solutions (those within the range) will have infinitely many solutions, while others will have no solutions.

(5) It is self-adjoint. Since we're dealing with Hilbert spaces, we have a self-dual space. So the inner product notation we've been using to denote evaluation is actually just the inner product now. That is, if $f \in L^2(\mu)$, then

$$f^* = \langle f, \cdot\rangle = \int f \cdot d\mu,$$

so that

$$\langle g, f^*\rangle = f^*(g) = \int fg d\mu = \langle g, f\rangle.$$

The adjoint then works nicely. Recall that the adjoint is defined to be the unique map T^* satisfying

$$\langle Tf, g\rangle = \langle f, T^* g\rangle.$$

Thus, we see

$$\langle Tf, g\rangle = \int (Tf)(s)g(s)d\mu(s) = \int \left(\int K(s,t)f(t)d\mu(t)\right) g(s)d\mu(s)$$
$$= \int g(s) \int K(s,t)f(t)d\mu(t)d\mu(s).$$

We now need to justify switching the order of integration. To do this is trickier than it seems. First, notice that $(Tf)(s) \in L^2(\mu)$, so by Cauchy-Schwarz we see that $g(Tf) \in L^1(\mu)$, since

$$\left|\int g(Tf)\right| \le \int |g||Tf| \le \left(\int |g|^2\right)\left(\int |Tf|^2\right) < \infty.$$

Define T^* by

$$(T^*g)(t) = \int K(s,t)g(s)d\mu(s).$$

Then we see that $(T^*g)f \in L^1(\mu)$ by a similar argument, and so switching the order of integration is fine. This gives

$$\langle Tf, g\rangle = \int\int K(s,t)f(t)g(s)d\mu(t)d\mu(s) = \int\int K(s,t)f(t)g(s)d\mu(s)d\mu(t) = \langle f, T^*g\rangle.$$

□

Problem 32 (Rudin 4.16, James)**.** Define

$$K(s,t) = \begin{cases} (1-s)t \text{ if } 0 \le t \le s \\ (1-t)s \text{ if } s \le t \le 1. \end{cases}$$

Define $T \in B(L^2(0,1))$ by

$$(Tf)(s) = \int_0^1 K(s,t)f(t)dt.$$

(1) Show that the eigenvalues of T are $(n\pi)^{-2}$, $n = 1, 2, \ldots$, that the corresponding eigenfunctions are $\sin(n\pi x)$ and that each eigenspace is one-dimensional.
(2) Show that the above eigenfunctions form an orthogonal base for $L^2(0,1)$.
(3) Suppose $g(t) = \sum c_n \sin(n\pi t)$. Discuss the equation $Tf - \lambda f = g$.
(4) Show that T is also a compact operator on $C([0,1])$.

Proof.

(1) Assume $\lambda \ne 0$, λ an eigenvalue. Then there is an f so that $(Tf) = \lambda f$. That is, for every $s \in [0,1]$ we have

$$(Tf)(s) = \int_0^1 K(s,t)f(t)dt = (1-s)\int_0^s tf(t)dt + s\int_s^1 (1-t)f(t)dt = \lambda f(s).$$

We see that taking the derivative gives

$$\lambda f'(s) = -\int_0^s tf(t)dt + (1-s)sf(s) + \int_s^1 (1-t)f(t)dt - s(1-s)f(s)$$
$$= \int_s^1 (1-t)f(t)dt - \int_0^s tf(t)dt.$$

Taking a second derivative gives

$$\lambda f''(s) = -(1-s)f(s) - sf(s) = -f(s).$$

So

$$\lambda f''(s) + f(s) = 0.$$

Note that

$$\lambda f(0) = (Tf)(0) = 0 \implies f(0) = 0,$$
$$\lambda f(1) = (Tf)(1) = 0 \implies f(1) = 0.$$

So we need to solve this second order differential equation. Examine the characteristic function:

$$\lambda r^2 + 1 = 0 \implies r^2 + \frac{1}{\lambda} = 0.$$

Assume $\lambda > 0$. This implies that the roots are complex; that is, the roots are

$$r = \pm\frac{i}{\sqrt{\lambda}}.$$

So the solutions are of the form

$$e^{i/\sqrt{\lambda}}, e^{-i/\sqrt{\lambda}}.$$

Using DeMoivre's formula, this implies

$$f(s) = C_1 \cos(s/\sqrt{\lambda}) + C_2 \sin(s/\sqrt{\lambda}).$$

Plugging in our initial conditions, we have

$$f(0) = C_1 = 0,$$
$$f(1) = C_2 \sin(1/\sqrt{\lambda}) = 0 \implies C_2 = 0 \text{ or } \lambda = (n\pi)^{-2} \text{ for } n \in \mathbb{N}.$$

We want to find the non-trivial solution, so this means that we have the eigenvalues are $(n\pi)^{-2}$ for $n \in \mathbb{N}$ and the eigenfunctions are

$$f(s) = \sin(n\pi s).$$

Notice this implies the eigenspace has dimension 1, since scalar multiples of these are the only solutions to this differential equation.

(2) We first show these are orthogonal. Assume $m, n \in \mathbb{Z}$ are distinct. Then we have

$$\int_0^1 \sin(n\pi x)\sin(m\pi x)dx = \frac{-m\cos(m\pi)\sin(n\pi) + n\sin(m\pi)\cos(n\pi)}{\pi(m^2 - n^2)} = 0.$$

We note we can turn this into an orthonormal set by just multiplying the components by $\sqrt{2}$. We now show completeness.

For all $f \in L^2((0,1))$, we can extend it (uniquely) to $L^2((-1,1))$ by taking an even extension, i.e.

$$\overline{f}(t) = \begin{cases} f(t) \text{ if } t \in [0,1] \\ -f(-t) \text{ if } t \in [-1,0]. \end{cases}$$

Recall that $L^2((-1,1))$ has an orthogonal basis given by $\{1, \cos(n\pi t), \sin(n\pi t)\}$ by Stone-Weierstrass. Now we see that

$$\int_{-1}^1 \overline{f}(t)\cos(n\pi t)dt = 0 \text{ for } n \geq 0,$$

and

$$\int_{-1}^1 \overline{f}(t)\sin(n\pi t)dt = 2\int_0^1 f(t)\sin(n\pi t)dt \text{ for } n \geq 1.$$

If

$$\int_0^1 f(t)\sin(n\pi t)dt = 0 \text{ for } n \geq 1,$$

then we have that the above integral is 0, implying that f is 0 since it evaluates to 0 on all of the orthogonal base. So we get that the set is complete, implying that it is an orthogonal base.

(3) Write

$$g(t) = \sum_{n=1}^{\infty} c_n \sin(n\pi t).$$

If g is 0, then the statement reduces to looking at

$$Tf = \lambda f,$$

for which there is only a (non-trivial) solution if λ is an eigenvalue. Assume now $g(t) = C\sin(n\pi t)$, where $C \neq 0$ is some constant and $n \geq 1$ an integer. We have that

$$Tf - \lambda f = C\sin(n\pi t).$$

Setting

$$f = (n\pi)^2 \sin(n\pi t),$$

we see

$$Tf = \sin(n\pi t),$$

so

$$Tf - \lambda f = \sin(n\pi t) - \lambda(n\pi)^2 \sin(n\pi t) = [1 - \lambda(n\pi)^2] \sin(n\pi t) = C\sin(n\pi t).$$

There is a solution if

$$1 - \lambda(n\pi)^2 = C \Leftrightarrow \lambda = (n\pi)^{-2}[1 - C].$$

The question now is whether this solution is unique. If λ is an eigenvalue, i.e. if

$$C = \frac{m^2 - n^2}{m^2}$$

for some $m \in \mathbb{Z}_{\geq 1}$, then the Fredholm alternative says that this is not a unique solution (scale things around). Otherwise, this is the unique solution.

For a finite sum, i.e. for

$$g(t) = \sum_{n=1}^{N} c_n \sin(n\pi t),$$

we can apply the prior strategy for each individual c_n and sum them together, giving us the existence of a solution as long as λ satisfies

$$\lambda = (n\pi)^{-2}[1 - c_n] \text{ for } 1 \leq n \leq N.$$

If this is the case, for each n we set $f_n = \sin(n\pi t)$ and we set

$$f = \sum_{n=1}^{N} f_n.$$

This tells us

$$Tf - \lambda f = \sum_{n=1}^{N} (Tf_n - \lambda f_n) = \sum_{n=1}^{N} c_n \sin(n\pi t).$$

Even if this strategy doesn't work, as long as λ is *not* an eigenvalue the Fredholm alternative tells us there is a unique solution. The question now is if λ is an eigenvalue, when do we have a solution for g?

Note by (2) we have an orthogonal base given by the sines. Write

$$f = \sum_{n=1}^{\infty} a_n \sin(n\pi t).$$

Notice

$$Tf = \sum_{n=1}^{\infty} a_n (n\pi)^{-2} \sin(n\pi t).$$

Since λ an eigenvalue, we know $\lambda = (\pi m)^{-2}$ for some $m \geq 1$ an integer. Hence we have

$$Tf - \lambda f = \sum_{n=1}^{\infty} a_n (n\pi)^{-2} \sin(n\pi t) - \sum_{n=1}^{\infty} a_n (m\pi)^{-2} \sin(n\pi t) = \sum_{n=1}^{\infty} c_n \sin(n\pi t)$$
$$\implies a_n((n\pi)^{-2} - (m\pi)^{-2}) = c_n \text{ for } n \geq 1.$$

We see there exists a solution for g so long as $c_m = 0$.

(4) We follow the hint in Rudin. Let $\{f_i\}$ be uniformly bounded. The goal is to show the (Tf_i) are equicontinuous. If we can show this, we invoke Arzela-Ascoli to deduce that the operator must be compact.

Note that (f_i) being uniformly bounded implies that $|f_i(x)| \leq M$ for all i and all $x \in [0, 1]$, $M < \infty$. Notice that for all i we have

$$|Tf_i(s) - Tf_i(r)| = \left| \int_0^1 K(s,t) f_i(t) dt - \int_0^1 K(r,t) f_i(t) dt \right|$$
$$= \left| (1-s) \int_0^s t f(t) dt + s \int_s^1 (1-t) f(t) dt - (1-r) \int_0^r t f(t) dt - r \int_r^1 (1-t) f(t) dt \right|$$
$$= \left| (1-s-1+r) \int_0^s t f(t) dt + (1-r) \int_0^s t f(t) dt + (s-r) \int_s^1 (1-t) f(t) dt \right.$$
$$\left. + r \int_s^1 (1-t) f(t) dt - (1-r) \int_0^r t f(t) dt - r \int_r^1 (1-t) f(t) dt \right|$$
$$= \left| (r-s) \int_0^s t f(t) dt + (1-r) \int_r^s t f(t) dt + (s-r) \int_s^1 (1-t) f(t) dt + r \int_s^r (1-t) f(t) dt \right|$$
$$\leq \frac{M}{2} |r-s| + M(1-r) \int_r^s t dt + \frac{M}{2} |s-r| + Mr \int_r^s (1-t) dt$$
$$= M|r-s| + \frac{M}{2} (r-s)(2r^2 + 2rs - 3r - s)$$
$$\leq M|r-s| \left(1 + r^2 + rs - \frac{3}{2} r - \frac{s}{2} \right).$$

Noting that $0 \leq r, s \leq 1$, we can maximize the above to get

$$|Tf_i(s) - Tf_i(r)| \leq \frac{3}{2} M |r-s|.$$

This is independent of the choice of i, so we can choose δ uniformly so that it applies for all i. Thus, we have equicontinuity.

□

1.5. **Chapter 10.**

Problem 33 (Rudin 10.1, James)**.** Use the identity

$$(xy)^n = x(yx)^{n-1}y$$

to prove that xy and yx always have the same spectral radius

Proof. Notice the identity follows by breaking it up; for $n \geq 1$, write $(xy)^n = xy \cdot \dots \cdot xy$ n times. Taking out the first x and the last y, we are left with $x(yx \cdot \dots \cdot yx)y$. We took out one copy of xy, so this leaves us with $n-1$ of these.

Recall from **Rudin 10.13** that the spectral radius of xy satisfies

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

Using the identity, we have

$$\begin{aligned} \rho(xy) = \lim_{n\to\infty} \|(xy)^n\|^{1/n} &= \lim_{n\to\infty} \|x(yx)^{n-1}y\|^{1/n} \\ &\leq \lim_{n\to\infty} \|x\|^{1/n}\|(yx)^{n-1}\|^{1/n}\|y\|^{1/n} \\ = \lim_{n\to\infty} \|(yx)^{n-1}\|^{1/n} = \lim_{n\to\infty} \left(\|(yx)^{n-1}\|^{1/(n-1)}\right)^{(n-1)/n} &= \rho(yx). \end{aligned}$$

A symmetric argument gives

$$\rho(yx) \leq \rho(xy),$$

so we have equality. □

Problem 34 (Rudin 10.2, James)**.**

(1) If x and xy are invertible in A, prove that y is invertible.
(2) If xy and yx are invertible in A, prove that x and y are invertible.
(3) Prove that it is possible to have $xy = e$ but $yx \neq e$.
(4) If $xy = e$ and $yx = z \neq e$, show that z is a non-trivial idempotent.

Proof.

(1) If x is invertible, there is an x^{-1}. Similarly, let $xy = z$, then there is a z^{-1}. So

$$(xy)z^{-1} = e \implies x(yz^{-1}) = e \implies yz^{-1} = x^{-1} \implies yz^{-1}x = e.$$

Notice as well

$$z^{-1}xy = z^{-1}(xy) = e.$$

So $y^{-1} = z^{-1}x$.

(2) Let $z = xy$, $q = yx$, then we see there is a z^{-1} and a q^{-1}. The goal is to prove that x is invertible (the same argument will show y is invertible). Notice

$$x = xqq^{-1} = x(yx)q^{-1} = (xy)xq^{-1} \implies z^{-1}x = xq^{-1}.$$

Now

$$e = z^{-1}z = z^{-1}(xy) = (z^{-1}x)y = (xq^{-1}y) = x(q^{-1}y).$$

Notice as well

$$q^{-1}yx = q^{-1}(yx) = e.$$

So $x^{-1} = q^{-1}y$.

(3) We follow the hint. Consider the left and right shifts S_r and S_f on some Banach space of functions, that is, set

$$(S_r f)(n) = f(n-1) \text{ if } n \geq 1,$$
$$(S_r f)(0) = 0,$$
$$(S_l f)(n) = f(n+1) \text{ if } n \geq 0.$$

Then $S_r \circ S_l(f)(n) = S_r(S_l(f))(n) = (S_l(f))(n-1) = f(n)$ as long as $n \geq 1$. At $n = 0$, $S_r(S_l(f))(0) = S_l(f)(0) = f(1)$, so we see it is not the identity. On the other hand, $S_l \circ S_r(f)(n) = S_l(S_r(f))(n) = S_r(f)(n+1) = f(n)$ for all $n \geq 0$, so it *is* the identity.

(4) Recall a non-trivial idempotent is an element satisfying $z^2 = z$ and $z \neq 0, z \neq e$. We see (using $xy = e$)

$$z^2 = (yx)(yx) = y(xy)x = yx.$$

Clearly $z \neq e$ by construction. We need to show $z \neq 0$. If $z = 0$, then this says $yx = 0$, so

$$y = y(xy) = (yx)y = 0 \cdot y = 0,$$

implying 0 is invertible, a contradiction.

□

Problem 35 (Rudin 10.3, James)**.** Prove that every finite dimensional A is isomorphic to an algebra of matrices.

Proof. Send basis elements to basis elements of $\mathbb{C}^n$ (since it's finite dimensional) and identify matrices correspondingly. □

Problem 36 (Rudin 10.5, James)**.** Let A_0 and A_1 be the algebras of all complex 2-by-2 matrices of the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \qquad \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix}.$$

Prove that every two-dimensional complex algebra A with unit e is isomorphic to one of these, and that A_0 is not isomorphic to A_1.

Proof. Since A is two dimensional, it admits a basis $\{e, a\}$ where e is such that $ae = ea = a$. Notice that we must have

$$a^2 = \alpha e + \beta a, \qquad \alpha, \beta \in \mathbb{C}.$$

We claim we can find another basis $\{e, b\}$ with $b^2 = \lambda e$ for some $\lambda \in \mathbb{C}$. If $\beta = 0$, we are done. If $\beta \neq 0$, notice that $\alpha \neq 0$ as well, for if $\alpha = 0$ we get

$$a^2 = \beta a \implies a(a - \beta e) = 0,$$

implying either $a = \lambda e$ for $\lambda \in \mathbb{C}$, contradicting the fact this is a basis. So

$$(a - \gamma e)^2 = a^2 - 2\gamma a + \gamma^2 e = \alpha e + \beta a - 2\gamma a + \gamma^2 e = (\alpha + \gamma^2)e + (\beta - 2\gamma)a.$$

So if we set $b = a - \beta/2e$, the above shows

$$b^2 = (\alpha + \beta^2/4)e.$$

Furthermore, we see $\{e, b\}$ forms a basis. So we can find a basis with $a^2 = \lambda e$ for some $\lambda \in \mathbb{C}$.

We now break this into two cases: $\lambda = 0$ (that is, $a^2 = 0$) or $\lambda \neq 0$. We claim in the case $\lambda \neq 0$ there is no element that squares to 0. To see this, suppose

$$c = \alpha e + \beta a, \quad c^2 = 0, \quad c \neq 0.$$

Then

$$c \cdot c = (\alpha e + \beta a) \cdot (\alpha e + \beta a) = \alpha^2 e + \alpha\beta a + \alpha\beta a + \beta^2 a^2.$$

Since $a^2 = \lambda e$, we can rewrite this as

$$c^2 = (\alpha^2 + \beta^2\lambda)e + 2\alpha\beta a.$$

We see $c^2 = 0$ implies that $\alpha + \beta^2\lambda = 0$ and $\alpha\beta = 0$. This is a contradiction, since the last equation implies either α or β is 0, the fact that $c \neq 0$ implies at least one must be non-zero, and the first equation implies that if one is zero, then so is the other. So if $c \neq 0$, then $c^2 \neq 0$.

Consider the case $\lambda = 0$. Define the map

$$\varphi : A \to A_1, \qquad \varphi(\alpha e + \beta a) = \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix}.$$

It follows that this is a linear map and $L(e) = \mathrm{Id}$, so we just need to check multiplication. That is,

$$\varphi((\alpha_1 e + \beta_1 a)(\alpha_2 e + \beta_2 a)) = \varphi(\alpha_1 e + \beta_1 a)L(\alpha_2 e + \beta_2 a).$$

Notice

$$(\alpha_1 e + \beta_1 a)(\alpha_2 e + \beta_2 a) = \alpha_1\alpha_2 e + (\alpha_1\beta_2 + \alpha_2\beta_1)a,$$

and

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ 0 & \alpha_1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_2 & \beta_2 \\ 0 & \alpha_2 \end{pmatrix} = \begin{pmatrix} \alpha_1\alpha_2 & \alpha_1\beta_2 + \alpha_2\beta_1 \\ 0 & \alpha_1\alpha_2 \end{pmatrix},$$

so this is indeed multiplicative, hence an algebra homomorphism. It's clearly surjective, and if $\varphi(\alpha e + \beta a) = 0$, then this implies that $\alpha, \beta = 0$, so this is injective. Thus we have an isomorphism.

Consider the case $\lambda \neq 0$. So $a^2 = \lambda e$. We wish to find a basis element b so that $b^2 = b$, $b \neq e$. All of the elements are of the form

$$b = \alpha e + \beta a,$$

so

$$b^2 = (\alpha e + \beta a)(\alpha e + \beta a) = (\alpha^2 + \lambda\beta^2)e + 2\alpha\beta a.$$

We want

$$\alpha^2 + \lambda\beta^2 = \alpha,$$

$$\beta = 2\alpha\beta.$$

This forces

$$\alpha = \frac{1}{2}, \qquad \beta = \pm\frac{1}{2\sqrt{\lambda}}.$$

Moreover, $\alpha, \beta \neq 0$, so this implies we can find $b, c \in A$ with $b^2 = b$, $c^2 = c$, $b \neq c$, $bc = cb = 0$, and $b + c = e$. Moreover, these form a basis for our algebra. We then construct our homomorphism as

$$\varphi : A \to A_0, \qquad \varphi(\alpha b + \beta c) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

Linearity is again clear, injectivity is clear, surjectivity is clear, $L(e) = 1$, and so it suffices to show multiplicativity. Notice

$$(\alpha b + \beta c)(\alpha b + \beta c) = \alpha^2 b + \alpha\beta cb + \alpha\beta bc + \beta^2 c = \alpha^2 b + \beta^2 c.$$

Notice

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} \alpha^2 & 0 \\ 0 & \beta^2 \end{pmatrix}.$$

So we have that every two-dimensional complex algebra A with unit e is isomorphic to one of these. By the remark on the case of $\lambda \neq 0$ having no (non-zero) element that squares to 0, we see that $A_0 \not\cong A_1$. So up to isomorphism there are only two options for a two dimensional complex Banach algebra. □

1.6. **Chapter 11.**

Problem 37 (Rudin 11.1, James)**.** Prove the following:

(1) No proper ideal of A contains any invertible element of A.
(2) If J is an ideal in a commutative Banach algebra A, then its closure $\overline{J}$ is also an ideal.

Proof.

(1) Suppose $J < A$ is a proper ideal, and let $y \in J$ be invertible. Then $y^{-1} \in A$, so $yy^{-1} = e \in J$. But $e \in J$ implies $ex = x \in J$ for all $x \in A$. Hence $J = A$, a contradiction.
(2) We need to show for all $y \in \overline{J}$, $x \in A$, we have $xy \in \overline{J}$. Since A a Banach algebra, we can express $y \in \overline{J}$ as the limit $y_n \to y$, where $y_n \in J$ for all n. Since J an ideal, $xy_n \in J$ for all n, and by continuity of multiplication this implies $xy \in \overline{J}$ as desired.

□

2. Penneys Solutions

Problem 38 (Penneys 3, James)**.** Consider

$$L^p[0,1] := \left\{ f : [0,1] \to \mathbb{C} \text{ measurable} : \left(\int |f|^p \right)^{1/p} < \infty \right\}$$

for $0 < p < 1$.

(1) Show that
$$d(f,g) = \int_0^1 |f(t) - g(t)|^p dt$$
is a well-defined translation-invariant metric on $L^p[0,1]$.
(2) Show that $L^p[0,1]$ with the metric given in (1) is a complete metric space.
(3) Prove that the only convex open subsets of $L^p[0,1]$ are $\varnothing$ and $L^p[0,1]$.
(4) Deduce that if (X,τ) is a locally convex topological vector space and $T : L^p[0,1] \to X$ is a continuous linear map, then $T = 0$.

Proof. (1) We first show that it is a metric. There are three properties we must establish.

(i) $d(f,g) \geq 0$: This follows since it's an integral of a non-negative function.
(ii) $d(f,g) = d(g,f)$: Notice that
$$d(f,g) = \int_0^1 |f(t) - g(t)|^p dt = \int_0^1 |-(g(t) - f(t))|^p dt$$
$$= \int_0^1 |g(t) - f(t)|^p dt = d(g,f).$$
(iii) $d(f,g) = 0$ if and only if $f = g$ almost everywhere: First, assume that $d(f,g) = 0$. Then we have
$$d(f,g) = \int_0^1 |f(t) - g(t)|^p dt = 0.$$
This can only happen if $|f(t) - g(t)|^p = 0$ almost everywhere (see **Folland Proposition 2.16**), which forces $|f(t) - g(t)| = 0$ almost everywhere, or $f = g$ almost everywhere. Next, if $f = g$ almost everywhere, then $|f(t) - g(t)| = 0$ almost everywhere, and so we see that $d(f,g) = 0$.
(iv) Finally, we need to establish the triangle inequality. This is the more interesting property. Let f, g, h be measurable functions from $[0,1]$ to $\mathbb{C}$. The goal is to show that
$$d(f,g) \leq d(f,h) + d(h,g).$$

Notice that

$$d(f,g) = \int_0^1 |f(t) - g(t)|^p dt.$$

We can add and subtract h to get

$$d(f,g) = \int_0^1 |f(t) - h(t) + h(t) - g(t)|^p dt.$$

We now use concavity. We claim for $a, b \geq 0$, we have

$$(a+b)^p \leq a^p + b^p.$$

If $b = 0$, then it follows. Assume $b > 0$. Then we can divide by b^p to get that the above is equivalent to

$$\left(1 + \frac{a}{b}\right)^p \leq 1 + \left(\frac{a}{b}\right)^p.$$

Let $t = a/b$. This can then be rewritten as

$$(1+t)^p \leq 1 + t^p \Leftrightarrow (1+t)^p - t^p \leq 1.$$

To prove the above claim, we differentiate with respect to t to get

$$p(1+t)^{p-1} - pt^{p-1} = p\left[(1+t)^{p-1} - t^{p-1}\right],$$

and this is less than 0 for $t > 0$. Hence, it is decreasing, so it is maximized at $t = 0$, which is 1, giving us the desired inequality.

Thus, we see that

$$d(f,g) = \int_0^1 |f(t) - h(t) + h(t) - g(t)|^p dt \leq \int_0^1 (|f(t) - h(t)|^p + |h(t) - g(t)|^p)\, dt$$
$$= d(f,h) + d(h,g).$$

Hence, it is a well-defined metric. We then need to show translation invariance. Let $f, g, h : [0,1] \to \mathbb{C}$ be measurable functions. We have that

$$d(f+h, g+h) = \int_0^1 |f(t) + h(t) - g(t) - h(t)|^p dt = \int_0^1 |f(t) - g(t)|^p dt = d(f,g).$$

So it is a well-defined translation-invariant metric on $L^p[0,1]$.

(2) (Proof from here) We now need to show it is complete. Let $(f_n) \subset L^p[0,1]$ be a Cauchy sequence; i.e., for all $\epsilon > 0$, there exists an N so that for $n, m \geq N$, we have

$$d(f_n, f_m) = \int_0^1 |f_n(t) - f_m(t)|^p dt < \epsilon.$$

Rudin claims this to be analogous to the $p \geq 1$ case. Choose a subsequence $(f_{n_j}) \subset L^p[0,1]$ of (f_n) so that

$$\int_0^1 |f_{n_j}(t) - f_{n_{j+1}}(t)|^p dt < \frac{1}{2^j}.$$

Let

$$g_n(t) = \sum_{j=1}^{n} |f_{n_{j+1}}(t) - f_{n_j}(t)|.$$

We first claim that

$$\int_0^1 |g_n(t)|^p dt \leq 1.$$

To see this, note that

$$\int_0^1 |g_n(t)|^p dt = \int_0^1 \left| \sum_{j=1}^n |f_{n_{j+1}}(t) - f_{n_j}(t)| \right|^p dt$$

$$\leq \int_0^1 \sum_{j=1}^n |f_{n_{j+1}}(t) - f_{n_j}(t)|^p dt = \sum_{j=1}^n \int_0^1 |f_{n_{j+1}}(t) - f_{n_j}(t)|^p dt$$

$$< \sum_{j=1}^n 2^{-j} \leq 1,$$

where here we used the claim from (1). Now, we write

$$g(t) = \sum_{j=1}^\infty |f_{n_{j+1}}(t) - f_{n_j}(t)|.$$

We claim that

$$\int_0^1 |g(t)|^p dt \leq 1.$$

To see this, note that

$$\int_0^1 |g(t)|^p dt = \int_0^1 \left| \sum_{j=1}^\infty |f_{n_{j+1}}(t) - f_{n_j}(t)| \right|^p dt$$

$$= \int_0^1 \lim_{n\to\infty} \left| \sum_{j=1}^n |f_{n_{j+1}}(t) - f_{n_j}(t)| \right|^p dt.$$

Recall that Fatou's Lemma says

$$\int \liminf g_n \leq \liminf \int g_n,$$

where g_n is a positive valued function. Hence, we have

$$\int_0^1 \lim_{n\to\infty} \left| \sum_{j=1}^n |f_{n_{j+1}}(t) - f_{n_j}(t)| \right|^p dt \leq \liminf_{n\to\infty} \int_0^1 \left| \sum_{j=1}^n |f_{n_{j+1}}(t) - f_{n_j}(t)| \right|^p dt$$

$$\leq \liminf_{n\to\infty} \int_0^1 \sum_{j=1}^n |f_{n_{j+1}}(t) - f_{n_j}(t)|^p dt$$

$$= \liminf_{n\to\infty} \sum_{j=1}^n \int_0^1 |f_{n_{j+1}}(t) - f_{n_j}(t)|^p dt$$

$$< \liminf_{n\to\infty} \sum_{j=1}^n 2^{-j} \leq 1.$$

Define the function

$$G(t) = f_{n_1}(t) + \sum_{j=1}^\infty (f_{n_{j+1}}(t) - f_{n_j}(t)).$$

Then we have that the function is defined almost everywhere, so let $F(t)$ be the function defined to be the value of the series if it converges and 0 otherwise. Note that

$$F(t) = \lim_{j\to\infty} f_{n_j}(t) \text{ almost everywhere.}$$

Fatou's Lemma then applies, so we see that for fixed $\epsilon > 0$, we can find N sufficiently large so that for $n \geq N$,

$$d(F, f_n) = \int_0^1 |F(t) - f_n(t)|^p dt \leq \liminf_{j\to\infty} \int_0^1 |f_{n_j}(t) - f_n(t)|^p dt < \epsilon^p.$$

So $F - f_n$ is in L^p, hence F in L^p, and $f_n \to F$ with respect to the L^p metric.

(3) We now need to show that the only open convex sets are $\varnothing$ and $L^p[0,1]$. A set $U \subset L^p[0,1]$ is *convex* if, for $0 \leq t \leq 1$, we have that

$$tU + (1-t)U \subset U.$$

We follow **Rudin 1.47**. Let $U \subset L^p[0,1]$ be open and convex. Since a shifting of convex sets is convex, assume without loss of generality that $0 \in U$. Then we have that there is some $\delta > 0$ so that $B_\delta(0) \subset U$, where

$$B_\delta(0) = \left\{ f \in L^p[0,1] : \int_0^1 |f(t)|^p dt < \delta \right\}.$$

Let $f \in L^p[0,1]$. Since $f \in L^p[0,1]$, we get

$$\int |f(t)|^p dt = M < \infty.$$

Since $p < 1$, we have that there exists a positive integer n so that

$$n^{p-1} \int_0^1 |f(t)|^p dt < \delta.$$

Now, by the continuity of the indefinite integral of $|f(t)|^p$, we have that there is a partition of $[0,1]$

$$0 = x_0 < x_1 < \cdots < x_n = 1$$

so that

$$\int_{x_{i-1}}^{x_i} |f(t)|^p dt = n^{-1} \int_0^1 |f(t)|^p dt.$$

Let

$$g_i(t) = \begin{cases} nf(t) \text{ if } x_{i-1} < t \leq x_i \\ 0 \text{ otherwise.} \end{cases}$$

Then we have that $g_i \in U$, since

$$\int_0^1 |g_i(t)|^p dt = \int_{x_{i-1}}^{x_i} |nf(t)|^p dt = n^p \int_{x_{i-1}}^{x_i} |f(t)|^p dt = n^{p-1} \int_0^1 |f(t)|^p dt < \delta,$$

so $g_i \in B_\delta(0) \subset U$. Now, we have that U is convex, so

$$f = \frac{1}{n}(g_1 + \cdots + g_n) \in U.$$

Thus, $U = L^p[0,1]$.

(4) Finally, the goal is to deduce that if $T : L^p[0,1] \to X$ is a continuous linear map, then $T = 0$. Since it's linear, we have that $T(0) = 0$. Next, let $U \subset X$ be an open set. Since X is locally convex, we get that there is an open convex set V so that $V \subset U$. Thus, we get that $T^{-1}(V) \subset T^{-1}(U)$, and since preimages of convex sets are convex, we get that $T^{-1}(V)$ is a convex open set, so $T^{-1}(V) = T^{-1}(U) = L^p[0,1]$. This applies to all open sets, so in particular any open neighborhood of 0, so T must be the 0 mapping.

□

Problem 39 (Penneys 4, James)**.** Let (X, τ) be a topological vector space. For $x \in X$, let $\mathcal{O}(x)$ denote the collection of open neighborhoods of x. Prove the following.

(1) Every open $U \in \mathcal{O}(0_X)$ is absorbing.
(2) If $U, V \subset X$ are open, then so is $U + V$.
(3) If $U \subset X$ is open, then so is the convex hull of U.
(4) Every convex $U \in \mathcal{O}(0_X)$ contains a balanced convex $V \in \mathcal{O}(0_X)$.

Proof.

(1) Recall a set $U \subset X$ is *absorbing* if, for every $x \in X$, there exists a t so that $x \in tU$. Fix $x \in X$ then we have that $\alpha \mapsto \alpha x$ is a continuous mapping, so in particular the collection $\{\alpha : \alpha x \in U\}$ is open, and it contains 0. So we can find n sufficiently large so that $(1/n)x \in U$, which implies that $x \in nU$ for n sufficiently large. Hence, U is absorbing.

(2) The shift of an open set by a point is still open (since translation is a homeomorphism), so

$$U + V = \bigcup_{x \in V} (U + x)$$

is open.

(3) The *convex hull* of $U \subset X$ is the smallest convex set containing U. Denote this by $\mathrm{Conv}(U)$. Let K be a convex set containing U. Note that K^{o} is convex, since for fixed $0 \leq t \leq 1$ we have

$$tK^{\mathrm{o}} + (1-t)K^{\mathrm{o}} \subset K$$

is open. Hence, if $\mathrm{Conv}(U)$ were not open, then we would have $\mathrm{Conv}(U)^{\mathrm{o}}$ is open and convex and contains U while being smaller than $\mathrm{Conv}(U)$, contradicting the minimality of $\mathrm{Conv}(U)$. Thus, $\mathrm{Conv}(U)$ must be open.

(4) This is **Rudin Theorem 1.14 (b)**. Suppose U a convex neighborhood of 0. Let

$$A := \bigcap_{|\alpha|=1} \alpha U.$$

By the continuity of scalar multiplication, there is a neighborhood V and a $\delta > 0$ such that $\alpha V \subset U$ for $|\alpha| < \delta$. Let W be the union of the αV. W is then a balanced neighborhood of 0, $W \subset U$. Since W is balanced, we see that $\alpha^{-1} W = W$ when $|\alpha| = 1$, so $W \subset \bigcap_{|\alpha|=1} \alpha U = A$. A is convex, being an intersection of convex sets, and so by prior work we have that A^{o} is convex. Recall A is balanced implies A^{o} is balanced, and we see that A is balanced since if $\alpha = r\beta$ is such that $|\alpha| \leq 1$, $|\beta| = 1$, and $0 \leq r \leq 1$, then

$$\alpha A = r\beta A = \bigcap_{|z|=1} r\beta z U = \bigcap_{|z|=1} rzU \subset \bigcap_{|z|=1} zU = A.$$

So A is balanced.

□

Problem 40 (Penneys 5, James)**.** Suppose $\varphi, \varphi_1, \ldots, \varphi_n$ are linear functionals on a vector space X. Prove that the following are equivalent:

(1) We have that

$$\varphi = \sum_{k=1}^{n} \alpha_k \varphi_k,$$

where $\alpha_i \in \mathbb{R}$.

(2) There is an $\alpha > 0$ so that for all $x \in X$,

$$|\varphi(x)| \leq \alpha \max\{|\varphi_k(x)| : 1 \leq k \leq n\}.$$

(3) We have that

$$\bigcap_{k=1}^{n} \ker(\varphi_k) \subset \ker(\varphi).$$

Proof. (1) $\implies$ (2) : For all $x \in X$, we have

$$|\varphi(x)| = \left| \sum_{k=1}^{n} \alpha_k \varphi_k(x) \right| \leq \sum_{k=1}^{n} |\alpha_k||\varphi_k(x)|$$
$$\leq \max\{|\varphi_k(x)| : 1 \leq k \leq n\} \left(\sum_{k=1}^{n} |\alpha_k| \right).$$

Setting

$$\alpha = \sum_{k=1}^{n} |\alpha_k|$$

gives us the desired result.

(2) $\implies$ (3) : If $x \in \bigcap_{k=1}^{n} \ker(\varphi_k)$, then $\varphi_k(x) = 0$ for $1 \leq k \leq n$. This implies that

$$|\varphi(x)| \leq \alpha \max\{|\varphi_k(x)| : 1 \leq k \leq n\} = 0,$$

so $x \in \ker(\varphi)$.

(3) $\implies$ (1) : (**Rudin Lemma 3.9**) Define the map $f : X \to \mathbb{R}^n$ by

$$f(x) = (\varphi_1(x), \ldots, \varphi_n(x)).$$

Notice that $\ker(f) = \bigcap_{k=1}^{n} \ker(\varphi_k) \subset \ker(\varphi)$ by (3), so we get a linear functional $g : \mathrm{Im}(f) \to \mathbb{R}$ by $g(f(x)) = \varphi(x)$. Now extend g to a linear function $G : \mathbb{R}^n \to \mathbb{R}$, defined in such a way so that $G(f(x)) = \varphi(x)$. To do so, we note that $\mathrm{Im}(f)$ is finite dimensional, so has a basis $\{u_n\}$. Since $X/\ker(f) \cong \mathrm{Im}(f) \subset \mathbb{R}^n$, we can view these as basis elements in $\mathbb{R}^n$, so on all basis elements which are not in R we set G to be 0. Then $G : \mathbb{R}^n \to \mathbb{R}$ is so that $g = G$ on $\mathrm{Im}(f)$.

Now, $G(x) = \sum_{k=1}^{n} \alpha_k \pi_k(x)$, where $\pi_k : \mathbb{R}^n \to \mathbb{R}$ is the kth component projection. Then

$$G(f(x)) = G(\varphi_1(x), \ldots, \varphi_n(x)) = \sum_{k=1}^{n} \alpha_k \varphi_k(x) = \varphi(x).$$

This holds for all x, so we have (1) holds. ☐

Problem 41 (Penneys 6, James)**.** Suppose X is a vector space and Y is a separating linear space of functionals on X. Endow X with the weak topology induced by Y. Prove that a linear functional φ on X is weakly continuous if and only if $\varphi \in Y$.

Proof. (**Rudin Theorem 3.10**) Recall that a family of functionals is said to be *separating* if for all $x \neq y$ in X, there is a $f \in Y$ so that $f(x) \neq f(y)$.

Recall that the *weak topology* on X induced by Y is the weakest topology so that for all $f \in Y$, f is continuous.

If $\varphi \in Y$, then clearly φ is weakly continuous. Suppose that φ is weakly continuous. Then there exists a neighborhood V around the origin so that

$$V \subset \{x \in X : |\varphi(x)| < 1\}.$$

In pariticular, the weak topology is generated by balls of the form

$$V(\{p_i\}_{i=1}^k, \epsilon) = \{x \in X : |p_i(x)| < \epsilon\}$$

where $p_i \in Y$, $\epsilon > 0$, so we have that there exists a neighborhood $V(\{p_i\}_{i=1}^k, \epsilon)$ with

$$V(\{p_i\}_{i=1}^k, \epsilon) \subset \varphi^{-1}(B_1(0)).$$

Notice that this implies that

$$|\varphi(x)| \leq \epsilon \max\{|p_i(x)| : 1 \leq i \leq k\}.$$

By **Penneys 5**, this implies that φ is a linear combination of elements from Y, so is in Y. □

Problem 42 (Penneys 36, James)**.** Let A be a unital Banach algebra. Suppose we have a norm convergent sequence $(a_n) \subset A$ with $a_n \to a$. Prove that for every open neighborhood U of $\sigma(a)$, there is an $N > 0$ such that $\sigma(a_n) \subset U$ for all $n > N$.

Proof. This follows by **Rudin 10.20**. The goal is to show that there is an $\epsilon > 0$ such that if $z \in B_\epsilon(a)$, then $\sigma(z) \subset U$. If we establish this, then since there is an N such that $a_n \in B_\epsilon(a)$ for all $n > N$ by norm convergence, we get that $\sigma(a_n) \subset U$ for all $n > N$.

To find this ϵ, we use the fact that $\|(\lambda e - a)^{-1}\|$ is continuous with domain the complement of U (by simply noting it is the composition of continuous functions). This norm tends to 0 as $\lambda \to \infty$, so it is uniformly bounded. That is, there exists $M < \infty$ with

$$\|(\lambda e - a)^{-1}\| < M$$

for all $\lambda \notin U$. If $y \in A$, $\|y\| < 1/M$, $\lambda \notin U$, then

$$\lambda e - (a + y) = (\lambda e - a)[e - (\lambda e - a)^{-1} y]$$

is invertible in A since $\|(\lambda e - a)^{-1}\| < 1$. This implies $\lambda \notin \sigma(x + y)$. Take $\epsilon = 1/M$ to get the desired result. □

Problem 43 (Penneys 37, James)**.** Let X be a Banach space and let $[a, b] \subset \mathbb{R}$ be a compact interval. Let $C := C([a, b], X)$ be the space of continuous functions $[a, b] \to X$, X equipped with the norm topology.

(1) Show that every $f \in C$ is uniformly continuous.
(2) Prove that C is a Banach space under the supremum norm:

$$\|f\|_\infty := \sup_{t \in [a,b]} \|f(t)\|.$$

Proof.

(1) This is the usual compact metric space argument.
(2) It is clearly a vector space (do the usual tricks, noting addition is pointwise and scaling a continuous function is continuous). We walk through the norm definition (though this too is the usual tricks).
   (a) It maps to $\mathbb{R}_{\geq 0}$ since $\|\cdot\|$ is a norm (here one might need to mention supremum property for positiveness but this is easy).
   (b) If $\|f\|_\infty = 0$, then this implies that

   $$0 \leq \|f(t)\| \leq 0 \text{ for all } t \in [a, b]$$

   by supremum properties. Hence $\|f(t)\| = 0$ for all t. Since $\|\cdot\|$ is a norm, this implies $f(t) = 0$ for all t, so f is the zero function.

(c) Notice

$$\|cf\|_\infty = \sup_{t\in[a,b]} \|cf(t)\| = \sup_{t\in[a,b]} |c|\|f(t)\| = |c|\|f\|_\infty.$$

(d) The triangle inequality follows from noting that the supremum and the norm satisfy the triangle inequality.

The next thing to check is completeness of the norm. Let $(f_n) \subset C$ be a Cauchy sequence. This implies for all $\epsilon > 0$, there is an N so that for $n, m > N$, we have

$$\|f_n - f_m\|_\infty < \epsilon.$$

Note this implies for each t we have

$$\|f_n(t) - f_m(t)\| < \epsilon$$

by the supremum. Now $(f_n(t)) \subset X$ is a Cauchy sequence with respect to the norm. Since X is Banach, we have $f_n(t) \to y \in X$. Define $f(t) = y$ to be a function $f : [a, b] \to X$. We need to show $f_n \to f$ in the supremum norm. Notice, however,

$$\|f_n - f\|_\infty = \sup_{t\in[a,b]} \|f_n(t) - f(t)\|.$$

For each $t \in [a, b]$, we see that $\|f_n(t) - f(t)\| \to 0$, forcing the supremum to go to 0 as well, as desired.

□

Problem 44 (Penneys 38, James)**.** Let X be a Banach space. In this problem, we show that the Riemann integral for continuous paths $\gamma : [a, b] \to X$ is well-defined and compatible with X^*. Throughout, fix a continuous path $\gamma : [a, b] \to X$.

(1) A *partition* of $[a, b]$ is a finite list

$$P = \{a = t_0 < t_1 < \cdots < t_n = b\}.$$

Consider the set of partitions, denoted by Γ. We can give a partial ordering $\leq$ on Γ by saying $P \leq Q$ iff $P \subset Q$ as sets. Show that partitions form a directed set under this partial ordering.

(2) A *tagged partition* of $[a, b]$ is a pair (P, u), where P is a partition of $[a, b]$ (defined in (1)) and $u = (u_i)_{i=1}^n \in [a, b]^n$ is such that $t_{i-1} \leq u_i \leq t_i$ for all $i = 1, \ldots, n$. A *preorder* on a set is a binary relation which is both reflexive and transitive. Show that the partial order given on partitions in (1) induces a preorder on tagged partitions.

(3) For a tagged partition (P, u), let

$$x_{(P,u)} = \sum_{i=1}^{n} \gamma(u_i)(t_i - t_{i-1}).$$

Show that $x_{(P,u)}$ is a norm convergent net in X.

(4) Define

$$\int_0^1 \gamma(t)dt = \lim x_{(P,u)}.$$

Prove that for every $\varphi \in X^*$,

$$\varphi\left(\int_0^1 \gamma(t)dt\right) = \int_0^1 \varphi(\gamma(t))dt,$$

where the right hand side is the Riemann integral of $\varphi \circ \gamma : [a, b] \to \mathbb{C}$.

(5) Show that

$$\left\|\int_a^b \gamma(t)dt\right\| \le \int_a^b \|\gamma(t)\|dt.$$

Deduce that

$$\int_a^b : C([a,b],X) \to X$$

is a bounded linear transformation.

Proof.

(1) We remark the obvious notion that $\subset$ is a partial ordering on a collection of sets, so furthermore is a preorder. A directed set is a set equipped with a preorder satisfying the property that any two elements have an upper bound. Notice that if $P, Q \in \Gamma$, then $P \cup Q \in \Gamma$, where union is defined as sets. The union of two finite sets gives us a finite set, so $P \cup Q$ is indeed a partition.

(2) The induced preorder is given by $(P,u) \le (Q,r)$ if $u \subset r$ as sets (i.e. there are more tags) and $P \subset Q$ as sets. Note that with this, $(P,u) \le (P,u)$ (so it is reflexive) and if $(P,u) \le (Q,r)$, $(Q,r) \le (T,s)$, then $(P,u) \le (T,s)$ (since $u \subset s$ and $P \subset T$).

(3) Since γ is continuous, we have that it is uniformly continuous (by the usual argument). Denote

$$\|P\| = \max\{\Delta_i : i = 1, \dots, n\}, \qquad \Delta_i = t_i - t_{i-1}.$$

To show it is a norm convergent net, we need to show it is a Cauchy net. So for every $\epsilon > 0$, there is a (P,u) with $(Q,r),(T,s) \ge (P,u)$ implies

$$\|x_{(Q,r)} - x_{(T,s)}\| < \epsilon.$$

Notice

$$\|x_{(Q,r)} - x_{(T,s)}\| = \left\|\sum_{i=1}^m \gamma(r_i)\Delta_i - \sum_{i=1}^n \gamma(s_i)\Delta_i\right\|.$$

Eliminating the things they have in common (due to (P,u)) and relabeling leaves us with an upper bound of

$$\sum_{j=1}^k |\gamma(w_j)|\Delta_j.$$

Since these are refinements of (P,u), and noting by uniform continuity $\|\gamma\|_\infty < M$, choosing $\|P\| < \epsilon/Mk$ forces $|\Delta_j| < \epsilon/Mk$, and so we get an upper bound of ϵ. Since the choice of ϵ was arbitrary, the net must be Cauchy, giving us it converges to something in norm. Label the something as

$$\int_a^b \gamma(t)dt.$$

(4) We have

$$\varphi\left(\int_a^b \gamma(t)dt\right) = \varphi(\lim x_{(P,u)}) = \int_a^b \varphi(\gamma(t))dt.$$

The first equality is by definition. Notice that for each $x_{(P,u)}$, we have

$$\varphi(x_{(P,u)}) = \varphi\left(\sum_{i=1}^n \gamma(u_i)\Delta_i\right) = \sum_{i=1}^n \varphi(\gamma(u_i))\Delta_i$$

by linearity. By continuity of φ, we can take the limit of tagged partitions of this quantity instead. Doing so gives us the Riemann integral by definition.

(5) We have

$$\int_a^b \gamma(t)dt = x \in X.$$

We invoke Hahn-Banach (**Rudin 3.3**) to find a linear functional $\varphi \in X^*$ with $\varphi(x) = \|x\|$, $\varphi(y) \le \|y\|$ for all $y \in X$. So

$$\left\|\int_a^b \gamma(t)dt\right\| = \varphi\left(\int_a^b \gamma(t)dt\right) = \int_a^b \varphi(\gamma(t))dt \le \int_a^b \|\gamma(t)\|dt \le \|\gamma\|_\infty \int_a^b dt = (b-a)\|\gamma\|_\infty.$$

The operator norm of the integral is then $(b-a)$, so the integral is a bounded linear operator.

□

Problem 45 (Penneys 39, Thomas)**.** Let A be a unital Banach algebra. Show that the holomorphic functional calculus satisfies the following properties.

(1) Suppose $a \in A$ and $K \subset \mathbb{C}$ is compact such that $\sigma(a) \subset K^o$. Show there is an $M_K > 0$ such that for any $f \in H(K^o)$ which has continuous extension to K,

$$\|f(a)\| \le M_K \|f\|_{C(K)}.$$

(2) Suppose $(a_n) \subset A$ is a norm convergent sequence with $a_n \to a$. Show that for all $f \in \mathcal{O}(\sigma(a))$, $f(a_n) \to f(a)$.

Proof.

(1) Using the notation of Rudin, we have

$$\|\widetilde{f}(a)\| \le M_K \|f\|_{C(K)},$$

where

$$\widetilde{f}(a) = \frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx$$

and Γ is any contour that surrounds $\sigma(x)$ in K^o. Taking the norm and invoking the triangle inequality, we get

$$\begin{aligned}\|\widetilde{f}(a)\| = \left\|\frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx\right\| &\le \frac{1}{2\pi}\int_\Gamma |f(x)|\|xe-a\|^{-1}dx \\ &\le \|f\|_{C(K)}\left(\frac{1}{2\pi}\int_\Gamma \|xe-a\|^{-1}dx\right).\end{aligned}$$

Recall (**Rudin 3.30**, the resolvent identity) that $(xe-a)^{-1}$ is strongly holomorphic off of $\sigma(a)$. In particular, it is continuous on Γ, so $\|(xe-a)^{-1}\|$ is continuous on Γ. Hence

$$0 < M_K = \frac{1}{2\pi}\int_\Gamma \|(xe-a)^{-1}\|dx < \infty.$$

This tells us

$$\|\widetilde{f}(a)\| \le \|f\|_{C(K)} M_K$$

as desired.

(2) Use Problem 36 to get that eventually $\widetilde{f}(a_n)$ is well-defined on some open set U which contains $\sigma(a)$. Now note

$$\begin{aligned}\|\widetilde{f}(a) - \widetilde{f}(a_n)\| &= \frac{1}{2\pi}\left\|\int_\Gamma f(x)[(xe-a)^{-1}(xe-a_n)^{-1}]dx\right\| \\ &\le \frac{\|f\|_\Gamma}{2\pi}\int_\Gamma \|(xe-a)^{-1} - (xe-a_n)^{-1}\|dx,\end{aligned}$$

where here $\|f\|_\Gamma = \sup_{x\in\Gamma} |f(x)|$. Since Γ continuous, $(xe-a)^{-1}$ is uniformly continuous on Γ. We can take n sufficiently large so that for some choice of $\epsilon > 0$

$$\|(xe-a)^{-1} - (xe-a_n)^{-1}\| < \epsilon \frac{2\pi}{\|f\|_\Gamma \ell(\Gamma)},$$

so

$$\|\widetilde{f}(a) - \widetilde{f}(a_n)\| < \epsilon.$$

The choice of $\epsilon > 0$ was arbitary, so we get $\widetilde{f}(a_n) \to \widetilde{f}(a)$.

□

Problem 46 (Penneys 40, James)**.** Let A be a unital Banach algebra, and let $a, p \in A$ be such that $ap = pa$ (that is, they commute).

(1) Show that for every $f \in \mathcal{O}(\sigma(a))$, $\widetilde{f}(a)p = p\widetilde{f}(a)$.
(2) From here on, assume p is an idempotent. Show that pAp is a unital Banach algebra.
(3) Prove that $\sigma_{pAp}(pa) \subset \sigma_A(a)$.
(4) Prove that for every $\widetilde{f} \in \mathcal{O}(\sigma_A(a))$, $\widetilde{f}(ap) = p\widetilde{f}$ when viewed in the image of the holomorphic functional calculus

$$\mathcal{O}(\sigma_{pAp}(pa)) \to pAp : f \mapsto \widetilde{f}(pa).$$

Proof.

(1) We see that

$$\widetilde{f}(a) = \frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx.$$

So

$$\widetilde{f}(a)p = \left(\frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx\right)p = \frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}p\,dx.$$

By commutativity of a and p, we see

$$(xe-a)p - p(xe-a) = (xe)p - ap - p(xe) + pa = 0,$$

so

$$(xe-a)p = p(xe-a) \Leftrightarrow p(xe-a)^{-1} = (xe-a)^{-1}p.$$

Using this and the above, we get

$$\widetilde{f}(a)p = p\left(\frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx\right) = p\widetilde{f}(a).$$

(2) Recall p is an idempotent if $p^2 = p$. Notice

$$pAp = \{pxp : x \in A\}.$$

We need to show this satisfies all of the familiar properties for a Banach algebra. We first check it is a vector space. Taking $pxp, pyp \in pAp$, we see

$$pxp + pyp = p(xp + yp) = p((x+y)p) = p(x+y)p \in pAp,$$

so it is closed under addition. If $0 \in A$ is the unit with respect to addition, then

$$p0p + pxp = p(0+x)p = pxp = pxp + p0p,$$

so $p0p \in pAp$ is still the unit with respect to addition. It is closed under inverses for a similar reason. If $\alpha \in \mathbb{C}$ a scalar, then

$$\alpha(pxp) = p(\alpha x)p \in pAp,$$

so it is closed under scaling as well. The compatibility with respect to scaling follows easily, so it is clear that it is a complex vector space from here.

We check now that it is a complex algebra. Take $pxp, pyp, pzp \in pAp$. We see

$$pxp(pyppzp) = pxp(pyp^2zp) = pxp(pypzp) = pxpypzp \in pAp,$$

$$(pxppyp)pzp = (pxpyp)pzp = pxpypzp \in pAp,$$

so

$$pxp(pyppzp) = (pxppyp)pzp.$$

Next,

$$(pxp + pyp)pzp = pxpzp + pypzp = (pxp)(pzp) + (pyp)(pzp),$$

$$pxp(pyp + pzp) = pxpyp + pxpzp = (pxp)(pyp) + (pxp)(pzp).$$

Finally, if $\alpha \in \mathbb{C}$,

$$\alpha((pxp)(pyp)) = \alpha(pxpyp) = \alpha pxpyp = (\alpha pxp)(pyp) = pxp(\alpha pyp),$$

thanks to the fact that A is a complex algebra. So pAp is a complex algebra.

Assume that $p \neq 0$ (the case $p = 0$ makes all of the above trivial anyways). We see that

$$\|p\| = \|pp\| \leq \|p\|^2 \implies 1 \leq \|p\|,$$

since $p \neq 0$. Define

$$\|pap\|_p = \|p\|\|a\|.$$

A rescaling of a norm is a norm. We check that this is still a Banach space. Let $(px_np) \subset pAp$ be Cauchy. For $\epsilon > 0$, we have that there is a N so that for $n, m > N$,

$$\|px_np - px_mp\|_p = \|p(x_n - x_m)p\|_p = \|p\|\|x_n - x_m\| < \epsilon,$$

so $(x_n) \subset A$ is a Cauchy sequence, hence $x_n \to x$. The claim is that $px_np \to pxp$. This follows, since

$$\|px_np - pxp\|_p = \|p(x_n - x)p\|_p = \|p\|\|x_n - x\| \to 0.$$

So pAp is a Banach space with respect to this norm.

We now check it is a Banach algebra. Notice

$$\|(pxp)(pyp)\|_p = \|pxpyp\|_p = \|p\|\|xpy\| \leq \|p\|\|x\|\|p\|\|y\| = \|pxp\|_p\|pyp\|_p.$$

Hence, pAp is a Banach algebra.

It is unital with unit $p = (pep)$, since

$$(pxp)(p) = pxp^2 = pxp.$$

(3) If $x \notin \sigma_A(a)$, we have $(xe - a)$ is invertible, so there is some $q \in A$ with $q(xe - a) = e$. We see

$$pqp(xe - pa) = (pqp)xe - (pqp)pa = p(qxe)p - p(qa)p = p[q(xe - a)]p = pep = p.$$

So $x \notin \sigma_{pAp}(pa)$. Thus, $\sigma_{pAp}(pa) \subset \sigma_A(a)$.

(4) Let's first take the fact on faith. That is, assume the following.

Fact. Suppose $U \subset \mathbb{C}$ open and $\sigma_A(a0 \subset U$. Suppose as well that $\Phi : H(U) \to A$ is a homomorphism satisfying the following:

- $\Phi(z \mapsto 1) = 1_A$,
- $\Phi(z \mapsto z) = a$,
- If $(f_n) \subset H(U)$ converges locally uniformly to f, then $\Phi(f_n) \to \Phi(f)$.

Then $\Phi(f) = \widetilde{f}(a)$ for all $f \in H(U)$. That is, Φ is the holomorphic functional calculus restricted to $H(U) \subset \mathcal{O}(\sigma_A(a))$.

Define $\Phi : H(U) \to pAp$ by $\Phi(f) = p(\widetilde{f}(a))$. We show that this is a homomorphism satisfying the criteria outlined in the fact.

- Let $f(z) = 1$. Then
$$\Phi(f) = p(\widetilde{f}(a)),$$
$$\widetilde{f}(a) = \frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx = \frac{1}{2\pi i}\int_\Gamma (xe-a)^{-1}dx = e,$$
where here we used **Rudin 10.23, 10.24**. So
$$\Phi(f) = p(\widetilde{f}(a)) = p,$$
which is the unit of pAp.
- Let $f(z) = z$, then
$$\widetilde{f}(a) = \frac{1}{2\pi i}\int_\Gamma f(x)(xe-a)^{-1}dx = \frac{1}{2\pi i}\int_\Gamma x(xe-a)^{-1}dx = a,$$
again by **Rudin 10.23, 10.24**. So
$$\Phi(f) = p(\widetilde{f}(a)) = pa.$$
- Let $f_n \to f$ uniformly locally. We can invoke **Rudin 10.27** to get
$$\widetilde{f_n}(a) \to \widetilde{f}(a).$$
So
$$\Phi(f_n) = p(\widetilde{f_n}(a)) \to p(\widetilde{f}(a)) = \Phi(f).$$
- We now show that Φ is an algebra homomorphism. The only interesting part is multiplication (the rest follows by linearity of the integral). Define $h(x) = f(x)g(x)$. The goal is to show $\Phi(h) = \Phi(f)\Phi(g)$. Notice by **Rudin 10.25**, we get $\widetilde{h} = \widetilde{f}\widetilde{g}$. So
$$\Phi(h) = p(\widetilde{h}(a)) = p(\widetilde{f}(a)\widetilde{g}(a)) = p^2\widetilde{f}(a)\widetilde{g}(a) = p\widetilde{f}(a)p\widetilde{g}(a) = \Phi(f)\Phi(g),$$
where here we used (1) for commutativity in the fourth equality.

So we see Φ satisfies the desired property and is a homomorphism. Hence, by the fact, Φ is the holomorphic functional calculus, implying $\widetilde{f}(pa) = \Phi(f) = p(\widetilde{f}(a))$.

We now need to show that the uniqueness of holomorphic functional calculus (HFC) proves the fact. We see that (according to the theorem in Penneys on page 12) the HFC is uniquely characterized by 3 properties. Since Φ is an algebra homomorphism (satisfying (1)) and Φ gives local convergence (satisfying (2)), we just need to show (3). But this follows by just noting that Φ sends polynomials to polynomials and then using Runge's theorem and the local uniform convergence property to deduce Φ agrees everywhere with the usual HFC.

□

Problem 47 (Penneys 41, James)**.** Let $A \in M_n(\mathbb{C})$. As best you can, describe $f(A)$, where $f \in \mathcal{O}(\sigma(A))$.

Proof. I'm following this.

Fix $A \in M_n(\mathbb{C})$. Recall we have
$$\psi : \mathbb{C}[z] \to M_n(\mathbb{C}), \qquad \psi(p) = p(A).$$
We have an associated minimal polynomial of A, m_A, which can be written as
$$m_A = \prod_{i=1}^k (z-\lambda_i)^{e_i},$$
where $\lambda_i \in \sigma(A)$ for $1 \le i \le k$. We see we have an induced homomorphism
$$\widehat{\psi} : \mathbb{C}[z]/(m_A) \to M_n(\mathbb{C}), \qquad \widehat{\psi}(p + (m_A)) = p(A).$$

By the uniqueness of HFC, if we can construct a "nice" unital algebra homomorphism $\Phi : \mathrm{Hol}(U) \to M_n(\mathbb{C})$, then we can describe what f looks like with this homomorphism.

Define a homomorphism

$$\eta : \mathbb{C}[z]/(m_A) \to \mathrm{Hol}(U)/m_A \cdot \mathrm{Hol}(U), \qquad \eta(p + (m_A)) = p|_U + m_A \cdot \mathrm{Hol}(U).$$

The goal is to show that this is an isomorphism. If we denote $\pi : \mathrm{Hol}(U) \to \mathrm{Hol}(U)/m_A \cdot \mathrm{Hol}(U)$ as the canonical surjection, then we can construct $\Phi : \mathrm{Hol}(U) \to M_n(\mathbb{C})$ by setting $\Phi = \widehat{\psi} \circ \eta^{-1} \circ \pi$.

So our first step here is showing η is an isomorphism. First, we show it is injective. Suppose $p|_U \in m_A \cdot \mathrm{Hol}(U)$; that is, there is some $f \in \mathrm{Hol}(U)$ with $p|_U = m_A \cdot f$. Then $p|_U$ has a zero of order e_j at λ_j for each j. By the uniqueness of power series representation, $m_j := (z - \lambda_j)^{e_j}$ divides p in $\mathbb{C}[z]$, which implies $m_A \mid p$, or $p = 0$ in $\mathbb{C}[z]/(m_A)$.

We now show surjectivity. Write

$$r_l := \frac{m_A}{m_l} = \prod_{j \neq l} m_j.$$

We see $\gcd(r_1, \ldots, r_k) = 1$. We can then find $q_j \in \mathbb{C}[z]$ so that

$$1 = \sum_{j=1}^{k} q_j r_j.$$

Take $f \in \mathrm{Hol}(U)$. The goal is to find p with $m_A \mid (f - p)$ in $\mathrm{Hol}(U)$. Around each λ_j, we can write

$$f(z) = \sum_{n=0}^{\infty} a_{jn}(z - \lambda_j)^n, \qquad a_{jn} := \frac{f^n(\lambda_j)}{n!}.$$

Set

$$p_j = \sum_{n=0}^{e_j - 1} a_{jn}(z - \lambda_j)^n, \qquad 1 \leq j \leq k,$$

and

$$p := \sum_{j=1}^{k} p_j q_j r_j.$$

We see $f - p_j$ has a zero of order e_j at λ_j, so $m_j \mid (f - p_j)$ in $\mathrm{Hol}(U)$. Thus

$$f - p = \sum_{j=1}^{k} (f - p_j) q_j r_j = m_A \sum_{j=1}^{k} \frac{f - p_j}{m_j} q_j \in m_A \cdot \mathrm{Hol}(U).$$

We have then

$$\Phi(f) = p(A), \qquad p(A) = \sum_{j=1}^{k} p_j(A) q_j(A) r_j(A) = \sum_{j=1}^{k} \sum_{n=0}^{e_j - 1} \frac{f^{(n)}(\lambda_j)}{n!} (A - \lambda_j)^n q_j(A) r_j(A).$$

The only property not clear from this is the local uniform convergence. However, local uniform convergence implies the derivatives converge uniformly locally, and so we just use the representation above and linearity of the limit to get our result. □

Problem 48 (Penneys 41, Thomas)**.** Let $A \in M_n(\mathbb{C})$. As best you can, describe $f(A)$, where $f \in \mathcal{O}(\sigma(A))$.

Thomas' Proof. Let $A \in M_n(\mathbb{C})$ and suppose $f \in \mathcal{O}(\sigma(A))$. First suppose A is an $n \times n$ Jordan block,

$$A = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & \cdots & \lambda \end{pmatrix}.$$

Since f is holomorphic on an open set containing λ, we can express it as[1]

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\lambda)}{k!}(z-\lambda)^k.$$

Using this, we can write the HFC as

$$\widetilde{f}(A) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\lambda)}{k!}(A-\lambda I)^k.$$

Notice that

$$A - \lambda I = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

so there exists N such that $(A-\lambda I)^{N-1} \neq 0$, $(A-\lambda I)^N = 0$. Hence,

$$\widetilde{f}(A) = \sum_{k=0}^{N-1} \frac{f^{(k)}(\lambda)}{k!}(A-\lambda I)^k.$$

We now know how the HFC acts on Jordan blocks. The next step is to show how it acts on a Jordan matrix. Suppose

$$A = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_r, \end{pmatrix}$$

where the J_i are Jordan blocks for $1 \leq i \leq r$. We can express A as

$$A = \bigoplus_{i=1}^{r} J_i.$$

If we determine how the HFC acts on direct sums, we will know how it acts on Jordan matrices. We check for just two matrices and use induction.

Suppose

$$A = B \oplus C.$$

Let $k \in \mathbb{Z}$. We can write

$$\begin{aligned}(A-\lambda I)^k = (B \oplus C - \lambda I)^k &= ((B-\lambda I) \oplus (C - \lambda I))^k \\ &= (B-\lambda I)^k \oplus (C-\lambda I)^k,\end{aligned}$$

[1]Implicitly throughout, we will take disjoint open sets around each of the $\lambda \in \sigma(A)$ when we express the HFC.

where the identity is taken over the appropriate subspace in each step. Writing out the HFC, we have

$$\widetilde{f}(A) = \frac{1}{2\pi i}\int_\Gamma \widetilde{f}(\lambda)(A-\lambda I)^{-1}d\lambda = \frac{1}{2\pi i}\int_\Gamma \widetilde{f}(\lambda)[(B-\lambda I)^{-1}\oplus (C-\lambda I)^{-1}]d\lambda$$
$$(*) = \frac{1}{2\pi i}\int_\Gamma \widetilde{f}(\lambda)(B-\lambda I)^{-1}d\lambda + \frac{1}{2\pi i}\int_\Gamma \widetilde{f}(\lambda)(C-\lambda I)^{-1}d\lambda$$
$$= \widetilde{f}(B) + \widetilde{f}(C).$$

Equality (*) follows from expanding definitions. That is, if we take $\Lambda \in (X\oplus Y)^*$, we can write $\Lambda = \Lambda_1 \oplus \Lambda_2$ for $\Lambda \in X^*$, $\Lambda_2 \in Y^*$, and for $x \in X$, $y \in Y$ we have

$$(\Lambda_1\oplus\Lambda_2)(x\oplus y) = \Lambda_1 x + \Lambda_2 y.$$

Putting everything together and inducting, we see that

$$\widetilde{f}(A) = \sum_{i=1}^r \widetilde{f}(J_i).$$

Recall that a matrix B is similar to A if there exists P with $B = PAP^{-1}$. Assuming B is similar to A, we have

$$(B-\lambda I)^k = (PAP^{-1}-\lambda I)^k = [P(A-\lambda I)P^{-1})]^k.$$

Letting $C = (A-\Lambda I)$, we see that

$$[PCP^{-1}]^k = \underbrace{(PCP^{-1})\cdots(PCP^{-1})}_{k \text{ times}} = PC^kP^{-1}.$$

Using this identity, we get

$$\widetilde{f}(B) = \sum_{k=0}^\infty \frac{f^{(k)}(\lambda)}{k!}(B-\lambda I)^k = \sum_{k=0}^\infty \frac{f^{(k)}(\lambda)}{k!}P(A-\lambda I)^kP^{-1}$$
$$= P\left[\sum_{k=0}^\infty \frac{f^{(k)}(\lambda)}{k!}(A-\lambda I)^k\right]P^{-1}$$
$$= P\widetilde{f}(A)P^{-1}.$$

We now put this all together. Every matrix is similar to a matrix in Jordan normal form, so for $A \in M_n(\mathbb{C})$ we have there is a P with $PCP^{-1} = A$, C in Jordan normal form. We write

$$C = \bigoplus_{i=1}^r J_i.$$

Using the the direct sum, Jordan block, and conjugation properties, we see that

$$\widetilde{f}(A) = P\left[\sum_{i=1}^r \widetilde{f}(J_i)\right]P^{-1} = \sum_{i=1}^r\sum_{k=0}^{N_i-1}\frac{f^{(k)}(\lambda_i)}{k!}P(J_i-\lambda_i I)^kP^{-1}$$

□

References

[1] Gerald Folland. *Real Analysis: Modern Techniques and Their Applications*. Second Edition.
[2] David Penneys. *Functional Analysis 7211 Exercises*. Link.
[3] Walter Rudin. *Functional Analysis*. Second Edition.