1. Normed vector spaces

In this section $\mathbb{F}$ stands for either $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{X}$ be a vector space over $\mathbb{F}$.

1.1. Definitions and preliminary results.

Definition 1.1. A normed vector space $\mathcal{X} = (\mathcal{X}, \|\cdot\|)$ consists of a vector space $\mathcal{X}$ over $\mathbb{F}$ together with a norm $\|\cdot\| : \mathcal{X} \to [0,\infty)$ (see definition 5.1 – which does not change with $\mathbb{R}$ replaced by $\mathbb{C}$). We often denote the normed vector space as $\mathcal{X}$, with the norm $\|\cdot\|$ implicit.

As we noted before, using the properties of a norm, it is straightforward to check that $d : \mathcal{X} \times \mathcal{X} \to [0,\infty)$ defined by

$$d(x,y) := \|x-y\|$$

is a metric on $\mathcal{X}$. The resulting topology is the *norm topology* and it is the *default* topology on $\mathcal{X}$.

Definition 1.2. A normed vector space $\mathcal{X}$ is a *Banach space* if it is complete (with its norm topology). □

Definition 1.3. Two norms $\|\cdot\|_1, \|\cdot\|_2$ on $\mathcal{X}$ are *equivalent* if there exist constants $C, c > 0$ such that

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

for all $x \in \mathcal{X}$. □

Remark 1.4. Equivalent norms determine the same topology on $\mathcal{X}$ and the same Cauchy sequences (Problem 1.4). In particular, it follows that if $\mathcal{X}$ is equipped with two equivalent norms $\|\cdot\|_1, \|\cdot\|_2$ then it is complete (a Banach space) in one norm if and only if it is complete in the other.

Equivalence of norms is an equivalence relation on the set of norms on $\mathcal{X}$. □

The next proposition is simple but fundamental; it says that the norm and the vector space operations are continuous in the norm topology.

Proposition 1.5 (Continuity of vector space operations)**.** *Let $\mathcal{X}$ be a normed vector space over $\mathbb{F}$.*

a) *If (x_n) converges to x in $\mathcal{X}$, then $(\|x_n\|)$ converges to $\|x\|$ in $\mathbb{R}$.*
b) *If (k_n) converges to k in $\mathbb{F}$ and (x_n) converges to x in $\mathcal{X}$, then $(k_n x_n)$ converges to kx in $\mathcal{X}$.*
c) *If (x_n) converges to x and (y_n) converges to y in $\mathcal{X}$, then $(x_n + y_n)$ converges to $x + y$ in $\mathcal{X}$.*

Proof. The proofs follow readily from the properties of the norm, and are left as exercises. □

The following proposition gives a convenient criterion for a normed vector space to be complete.

Definition 1.6. Given a sequence (x_n) from a normed vector space $\mathcal{X}$, the expression $\sum_{n=1\infty} x_n$ denotes the sequence $(s_N = \sum_{n=1}^N x_n)$, called the sequence of *partial sums* of the series. The *series converges* if the sequence of partial sums converges in the norm topology. In this case we use $\sum_{n=1\infty} x_n$ to also denote the limit of this sequence and call it the sum.

Explicitly, the series $\sum_{n=1}^\infty x_n$ converges means there is an $x \in \mathcal{X}$ such that for each $\epsilon > 0$ there is an N such that $\|s_n - x\| < \epsilon$ for all $n \geq N$.

The series $\sum_{n=1}^\infty x_n$ *converges absolutely* if the series $\sum_{n=1}^\infty \|x_n\|$ converges (in the normed vector space $(\mathbb{R}, |\cdot|)$). □

Proposition 1.7. *A normed space $(\mathcal{X}, \|\cdot\|)$ is complete if and only if every absolutely convergent series in $\mathcal{X}$ is convergent.*

Before proving the Proposition we collect two lemmas. A definition is needed for the first.

Definition 1.8. *A sequence (y_k) from a normed vector space $\mathcal{X}$ is* super-cauhcy *if the series $\sum_{k=1}^\infty (y_{k+1} - y_k)$ converges absolutely.*

Lemma 1.9. *If (x_n) is a Cauchy sequence from a normed vector space $\mathcal{X}$, then there is a subsequence (y_k) of (x_n) that is super-cauchy.*

Proof. With $\epsilon = \frac{1}{2}$, there exists an N_1 such that $\|x_n - x_m\| < \frac{1}{2}$ for all $m, n \geq N_1$ since (x_n) is Cauchy. Assuming $N_1 < N_2 < \cdots < N_k$ have been chosen so that $\|x_n - x_m\| < \frac{1}{2^j}$ for $1 \leq j \leq k$, there is an $N_{k+1} < N_k$ such that $\|x_n - x_m\| < \frac{1}{2^{k+1}}$ since (x_n) is Cauchy. Hence by recursion we have constructed a (strictly)increasing sequence of integers N_k such that $\|x_n - x_m\| < \frac{1}{2^k}$ for all $m, n \geq N_k$. Set $y_k = x_{N_k}$ and note that $\|y_{k+1} - y_k\| < \frac{1}{2^k}$, from which it follows that (y_k) is a super-cauchy subsequence of (x_n). □

The proof will also use the following standard lemma from advanced calculus.

Lemma 1.10. *If (x_n) is a Cauchy sequence from a metric space (X, d) and if (x_n) has a subsequence (y_k) that converges to some x, then (x_n) converges to x.*

Proof of Proposition 1.7. First suppose $\mathcal{X}$ is complete and $\sum_{n=1}^\infty x_n$ is absolutely convergent. Write $s_N = \sum_{n=1}^N x_n$ for the N^{th} partial sum and let $\epsilon > 0$ be given. Since $\sum_{n=1}^\infty \|x_n\|$ is convergent, there exists an L such that $\sum_{n=L}^\infty \|x_n\| < \epsilon$. If $N > M \geq L$, then

$$\|s_N - s_M\| = \left\| \sum_{n=M+1}^{N} x_n \right\| \le \sum_{n=M+1}^{N} \|x_n\| < \epsilon.$$

Thus the sequence (s_N) is Cauchy in $\mathcal{X}$, hence convergent by the completeness hypothesis.

Conversely, suppose every absolutely convergent series in $\mathcal{X}$ is convergent and that (x_n) is given Cauchy sequence from X. By Lemma 1.9 there is a super-cauchy subsequence (y_k) of (x_n). Since (y_k) is super-cauchy, the series $\sum_{k=1}^{\infty}(y_{k+1} - y_k)$ is absolutely convergent and hence, by hypothesis, convergent in $\mathcal{X}$. Thus there is an $z \in \mathcal{X}$ such that the sequence of partial sums

$$\sum_{k=1}^{n}(y_{k+1} - y_k) = y_{n+1} - y_1$$

converges to z. Rearranging, $(x_{N_{n+1}} = y_{n+1})$ converges to $x = z + y_1$. Hence (x_n) is Cauchy and has a convergent subsequence. Thus (x_n) converges (to x) by Lemma 1.10. $\square$

1.2. **Examples.**

1.2.1. *Euclidean space.* Observe that the Euclidean norm on the complex vector space $\mathbb{C}^n$ agrees with the Euclidean norm on the real vector space $\mathbb{R}^{2n}$ (via that natural real linear isomomorphism $\mathbb{R}^2 \to \mathbb{C}$ sending (x, y) to $x + iy$). Thus, $\mathbb{F}^n$ with the usual Euclidean norm $\|(x_1, \ldots x_n)\| = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2}$ is a Banach space.

The vector space $\mathbb{F}^n$ can also be equipped with the ℓ^p*-norms*

$$\|(x_1, \ldots x_n)\|_p := \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}$$

for $1 \le p < \infty$, and the ℓ^∞*-norm*

$$\|(x_1, \ldots x_n)\|_\infty := \max(|x_1|, \ldots |x_n|).$$

For $1 \le p < \infty$ and $p \ne 2$, it is not immediately obvious that $\|\cdot\|_p$ defines a norm. We will prove this assertion later. It is not too hard to show that all of the ℓ^p norms $(1 \le p \le \infty)$ are equivalent on $\mathbb{F}^n$ (though the constants c, C depend on the dimension n). For instance, for $n \in \mathbb{N}$,

$$\|x\|_\infty \le \|x\|_2 \le \|x\|_1 \le n\|x\|_\infty.$$

The first and third inequalities are evident. For the middle inequality, observe

$$(\|x\|_1)^2 = \sum_{j,k=1}^{n} |x_j|\,|x_k| \ge \sum_{j=1}^{n} |x_j|^1 = \|x\|_2^2.$$

Given a normed vector space $\mathcal{X} = (\mathcal{X}, \|\cdot\|)$, denote its closed *unit ball* by

$$\mathcal{X}_1 = \{x \in \mathcal{X} : \|x\| \le 1\}.$$

It is instructive to sketch the closed unit ball in $\mathbb{R}^2$ with the three norms above.

It turns out that *any* two norms on a finite-dimensional vector space are equivalent. As a corollary, every finite-dimensional normed space is a Banach space. See Problem 1.5.

Lemma 1.11. *If $\|\cdot\|_1$ and $\|\cdot\|_2$ are norms on $\mathcal{X}$ and there is a constant $C > 0$ such that $\|x\|_1 \leq C\|x\|_2$ for all $x \in \mathcal{X}$, then the mapping $\iota : (\mathcal{X}, \|\cdot\|_2) \to (\mathcal{X}, \|\cdot\|_1)$ is (uniformly) continuous.*

Proof. For $x, y \in \mathcal{X}$, we have $\|\iota(x) - \iota(y)\|_1 = \|\iota(x-y)\|_1 = \|x-y\|_1 \leq C\|x-y\|_2$. □

Proposition 1.12. *If $\|x\|$ is a norm on $\mathbb{R}^n$, then $\|x\|$ is equivalent to the Euclidean norm $\|\cdot\|_2$.*

Sketch of proof. Let $\{e_1, \dots, e_n\}$ denote the usual basis for $\mathbb{R}^n$. Given $x = \sum a_j e_j \in \mathbb{R}^n$,

$$\|x\| \leq \sum |a_j|\,\|e_j\| = \sum |a_j|\|e_j\| \leq M\,\|x\|_1 \leq n\,M\,\|x\|_2,$$

where $M = \max\{\|e_1\|, \dots, \|e_n\|\}$. It now follows that the map $\iota : (\mathbb{R}^n, \|\cdot\|_2) \to (\mathbb{R}^n, \|\cdot\|)$ is continuous and therefore so is the map $f : (\mathbb{R}^n, \|\cdot\|_2) \to [0, \infty)$ defined by $f(x) = \|x\|$. Since

$$S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$$

(the unit sphere) is compact in $\mathbb{R}^n$, by the Extreme Value Theorem, f attains its infimum; that is, there is a point $p \in S^{n-1}$ such that $f(p) \leq f(x)$ for all $x \in S^{n-1}$. But $f(p) = \|p\| > 0$ since $p \neq 0$. Let $c = f(p) = \|p\|$. We conclude that if $\|x\|_2 = 1$ then $\|x\| \geq c\|x\|_2$, from which it follows by homogeneity that $\|x\| \geq c\|x\|_2$ for all $x \in \mathbb{R}^n$. □

Corollary 1.13. *All norms on a finite dimensional vector space are equivalent. Further, if V is a finite dimensional normed vector space, then V_1 is compact and V is a Banach space.*

Proof. Suppose V is a normed vector space of dimension n and let $\{v_1, \dots, v_n\}$ denote a basis for V. The function $\|\cdot\|' : V \to [0, \infty)$ defined by

$$\|v\|' = \|\sum a_j v_j\|' = \sum |a_j|$$

is easily seen to be a norm.

Now let $\|\cdot\|$ be a given norm on V. This norm induces a norm $\|\cdot\|_*$ on $\mathbb{R}^n$ given by

$$\|\sum a_j e_j\|_* = \|\sum a_j v_j\|.$$

Since all norms in $\mathbb{R}^n$ are equivalent, the norm $\|\cdot\|_*$ is equivalent to the norm $\|\cdot\|_1$. Hence there exist constants $0 < c < C$ such that

$$c\|v\|' = c\sum |a_j| = c\|\sum a_j e_j\|_1 \leq \|\sum a_j e_j\|_* \leq C\|\sum a_j e_j\|_1 = C\sum |a_j| = C\|v\|'.$$

Thus, as $\|\sum a_j e_j\|_* = \|\sum a_j v_j\|$,

$$c\|v\|' \leq \|v\| \leq C\|v\|'$$

for all $v \in V$. Thus all norms on V are equivalent.

Further, by definition, $f : (V, \|\cdot\|) \to (\mathbb{R}^n, \|\cdot\|_*)$ is bijective and isometric. Thus, f^{-1} is continuous, $f^{-1}(S)$ where S is the unit ball in $(\mathbb{R}^n, \|\cdot\|_*$, is the unit ball in $(V, \|\cdot\|)$ and is compact as its the continuous image of a compact set. It is now routine to pass from compactness of the unit ball in $(V, \|\cdot\|)$ to completeness of $(V, \|\cdot\|)$. □

1.2.2. *The Banach space of bounded functions.* If V is a vector space over $\mathbb{F}$ and $\varnothing \neq T$ is a set, then $F(T, V)$, the set of functions $f : T \to V$ is a vector space over $\mathbb{F}$ under pointwise operations; e.g., if $f, g \in F(T, V)$ then $f + g : T \to V$, is the function defined by $(f + g)(t) = f(t) + g(t)$.

Definition 1.14. A subset R of a normed vector space $\mathcal{X}$ is bounded if there is a C such that $\|x\| \leq C$ for all $x \in R$; that is, $R \subseteq C\mathcal{X}_1$.

A function $f : T \to \mathcal{X}$ is bounded if $f(T) \subseteq \mathcal{X}$ is bounded.

Let $F_b(T, \mathcal{X})$ denote the vector space (subspace of $F(T, \mathcal{X})$) of bounded functions $f : T \to \mathcal{X}$.

Remark 1.15. The function $\|\cdot\|_\infty : F_b(T, \mathcal{X}) \to [0, \infty)$ defined by

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

is a norm on $F_b(T, \mathcal{X})$ as you should verify. Let d_∞ denote the resulting metric: $d_\infty(f, g) = \|f - g\|_\infty$.

Note that convergence of a sequence in the metric space $(F_b(T, \mathcal{X}), d_\infty)$ is uniform convergence; in particular, a sequence is Cauchy in $F_b(T, \mathcal{X})$ if and only if it is uniformly Cauchy. (Exercise.)

Proposition 1.16. *If $\mathcal{X}$ is a Banach space, then $F_b(T, \mathcal{X})$ is also Banach space.*

Proof. We are to show $F_b(T, \mathcal{X})$ is complete, assuming $\mathcal{X}$ is complete. Accordingly, suppose (f_n) is a Cauchy sequence from $F_b(T, \mathcal{X})$ and $\mathcal{X}$ is complete. In particular, given $\epsilon > 0$ there is an N such that $d_\infty(f_n, f_m) = \sup\{\|f_n(t) - f_m(t)\| : t \in T\} < \epsilon$. It follows that, for each $s \in T$, the sequence $(f_n(s))$ is a Cauchy in $\mathcal{X}$ and hence converges to some $x \in \mathcal{X}$. Define $f : T \to \mathcal{X}$ by $f(s) = x$. It remains to see that f is bounded and (f_n) converges to f.

Since Cauchy sequences are bounded and (f_n) is Cauchy in the metric space $F_b(T, \mathcal{X})$, there is a C such that

$$\sup\{\|f_n(t)\| : t \in T\} = d_\infty(f_n, 0) \leq C$$

for all n. It follows from Proposition 1.5 that $(\|f_n(t)\|)_n$ converges to $|f(t)$ and hence $\|f(t)\| \leq C$ for all $t \in T$. Thus f is bounded; that is $f \in F_b(T, \mathcal{X})$.

It only remains to show that (f_n) converges to f in $F_b(T, \mathcal{X})$. To do so let $\epsilon > 0$ be given. There is an N such that if $m, n \geq N$, then $\|f_n(t) - f_m(t)\| < \epsilon$ for all $t \in T$.

Given $s \in T$, there is an $M \geq N$ such that $\|f_m(s) - f(s)\| < \epsilon$ for all $m \geq N$. Since, $(f_m(s) - f_n(s))_m$ converges (with m) to $(f(s) - f_n(s))$ in $\mathcal{X}$, another application of Proposition 1.5 gives $(\|f_m(s) - f_n(s)\|)_m$ converges to $\|f(s) - f_n(s)\|$. Thus

$$\|f(s) - f_n(s)\| \leq \epsilon,$$

for all $s \in T$. Hence $d_\infty(f, f_n) = \|f - f_n\| \leq \epsilon$ and the proof is complete. □

There are important Banach spaces of continuous functions. Before going further, we remind the reader of the following result from advanced calculus.

Theorem 1.17. *Suppose X, Y are metric spaces, (f_n) is a sequence $f_n : X \to Y$ and $x \in X$. If each f_n is continuous at x and if (f_n) converges uniformly to f, then f is continuous at x. Hence if each f_n is continuous, then so is f.*

Proof. Let x and $\epsilon > 0$ be given. Choose N such that if $n \geq N$ and $y \in X$, then $d_Y(f_n(y), f(y)) < \epsilon$. Since f_N is continuous at x, there is a $\delta > 0$ such that if $d_X(x, y) < \delta$, then $d_Y(f_N(x), f_N(y)) < \epsilon$. Thus, if $d_X(x, y) < \delta$, then

$$\begin{aligned} d_Y(f(x), f(y)) &\leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(y)) + d_Y(f_N(y), f(y)) \\ &< 3\epsilon, \end{aligned}$$

proving the theorem. □

Given a normed vector space $\mathcal{Y}$, let $C_b(X, \mathcal{Y})$ denote the subspace of $F_b(X, \mathcal{Y})$ consisting of continuous functions. Since uniform convergence is the same as convergence in the normed vector space $(F_b(X, \mathcal{Y}), d_\infty)$, by Theorem 1.17, $C_b(X, \mathcal{Y})$ is a closed subspace of $F_b(X, \mathcal{Y})$. In particular, in the case $\mathcal{Y}$ is a Banach space, so is $C_b(X, \mathcal{Y})$.

When X be a compact metric space, let $C(X) = C(X, \mathbb{F})$ denote the set of continuous functions $f : X \to \mathbb{F}$. Thus $C(X)$ is a subspace of $F_b(X, \mathbb{F})$ and we endow $C(X)$ with the norm it inherits from $F_b(X, \mathbb{F})$. Since $\mathbb{F}$ is complete, $C(X)$ is a Banach space. Of course here we could replace $\mathbb{F}$ by a Banach space $\mathcal{X}$ and obtain the analogous conclusion for the space $C(X, \mathcal{X})$.

Now let X be a locally compact metric space. In this case, a function $f : X \to \mathbb{F}$ *vanishes at infinity* if for every $\epsilon > 0$, there exists a compact set $K \subseteq X$ such that $\sup_{x \notin K} |f(x)| < \epsilon$. Let $C_0(X)$ denote the subspace of $F_b(X, \mathbb{F})$ consisting of continuous functions $f : X \to \mathbb{F}$ that vanish at infinity. Then $C_0(X)$ is a normed vector space with the norm it inherits from $C(X)$ (equivalently $F_b(X, \mathbb{F})$. It is routine to check that $C_0(X)$ is complete.

1.2.3. *L^1 spaces over $\mathbb{R}$.* Let $(X, \mathscr{M}, \mu)$ be a measure space and let $L^1(\mu)$ denote the (real) vector space of (real-valued) absolutely integrable functions on X. The function $\|\cdot\|_1 : L^1(\mu) \to [0, \infty)$ defined by

$$\|f\|_1 := \int_X |f| \, dm$$

for $f \in L^1(\mu)$ is a norm on $L^1(\mu)$, provided we agree to identify f and g when $f = g$ a.e. (Indeed the chief motivation for making this identification is that it makes $\|\cdot\|_1$ into a norm.

Proposition 1.18. *The real vector space $L^1(\mu)$ is a Banach space.*

We will construct a complex vector space analog of $L^1(\mu)$ a bit later.

Proof. It suffices to verify the hypotheses of Proposition 1.7. Accordingly suppose $\sum_{n=1}^\infty f_n$ is absolutely convergent (so that $\sum_{n=1}^\infty \|f_n\|_1 < \infty$). By Tonelli's summation theorem,

$$\int \sum_{n=1}^\infty |f_n| \, dm = \sum_{n=1}^\infty \int |f_n| \, dm = \sum_{n=1}^\infty \|f_n\|_1 < \infty.$$

Thus the function $g := \sum_{n=1}^\infty |f_n|$ belongs to L^1 and is thus finite m-a.e. In particular the sequence of partial sums $s_N = \sum_{n=1}^N f_n$ is a sequence of measurable functions with $|s_N| \le g$ that converges pointwise a.e. to a measurable function f. Hence by the DCT and its corollary, $f \in L^1$ and the partial sums $(s_N)_N$ converge to f in L^1. □

1.2.4. *Complex $L^1(\mu)$ spaces.* In this subsection we describe the extension of $L^1(\mu)$ to a complex vector space of complex valued functions (equivalence classes of functions).

Again we work on a fixed measure space $(X, \mathscr{M}, \mu)$. As a topological space, $\mathbb{C}$ and $\mathbb{R}^2$, are the same. A function $f : X \to \mathbb{C} = \mathbb{R}^2$ is *measurable* if and only if it is $\mathscr{M} - \mathscr{B}_2$ measurable. Measurability of f can also be described in terms of the real and imaginary parts of f.

Proposition 1.19. *Suppose $(X, \mathscr{M})$ is a measurable space and $f : X \to \mathbb{C}$. Writing $f : X \to \mathbb{C}$ as $f = u + iv$, where $u, v : X \to \mathbb{R}$, the function f is measurable if and only if both u and v are.*

Moreover, if f is measurable, then so is $|f| : X \to [0, \infty)$.

We begin with the following elementary lemma whose proof is left to the reader.

Lemma 1.20. *Suppose $(X, \mathscr{M})$ is a measure space and Y and Z are topological spaces. If $f : X \to Y$ is $\mathscr{M} - \mathscr{B}_Y$ measurable and $g : Y \to Z$ is $\mathscr{B}_Y - \mathscr{B}_Z$ measurable, then $g \circ f$ is $\mathscr{M} - \mathscr{B}_Z$ measurable. In particulr, the result holds if g is continuous.*

Sketch of proof of Proposition 1.19. The Borel σ-algebra $\mathscr{B}_2$ is generated by open rectangles; that is, a set $U \subseteq \mathbb{C}$ is open if and only if it is a countable union of open rectangles (with rational vertices even). For an open rectagle $I = J \times K = (a, b) \times (c, d)$ observe that

$$f^{-1}(J) = u^{-1}(J) \cap v^{-1}(K).$$

Thus, if u and v are measurable, then $f^{-1}(J) \in \mathscr{M}$. Consequently, f is measurable. Hence if u, v are both measurable, then so is f.

Now suppose f is measurable. In this case

$$\mathscr{M} \ni f^{-1}((t,\infty)\times\mathbb{R}) = \{u>t\}.$$

Since the sets $\{(t,\infty) : t\}$ generate $\mathscr{B}_1$, u is measurable. By symmetry v is measurable.

To prove the second statement, since f is measurable and $g:\mathbb{C}\to[0,\infty)$ defined by $g(z)=|z|$ is continuous, the function $g\circ f=|f|$ is measurable by Lemma 1.20. □

Definition 1.21. A measurable $f: X\to\mathbb{C}$ is *integrable* (or *absolutely integrable*) if $|f|$ is integrable.

Remark 1.22. From the inequalities

$$|\mathrm{Re} f|,\ |\mathrm{Im} f| \le |f| \le |\mathrm{Re} f| + |\mathrm{Im} f|$$

it follows that $f: X\to\mathbb{C}$ is (absolutely) integrable if and only if Ref and Imf are.

Definition 1.23. If f is complex-valued and absolutely integrable (that is, f is measurable and $|f|$ is integrable), we define the integal of f by

$$\int f = \int \mathrm{Re} f + i\int \mathrm{Im} f.$$

We also write $\|f\|_1 := \int_X |f|\,d\mu$ in the complex case. Finally, we write $L^1 = L^1(\mu)$ to denote the set of absolutely integrable complex-valued functions on X.

Generally, we leave it to context to indicate if we are considering the real or complex version of L^1; but for the following theorem we temporarily adopt the notation $L^1_{\mathbb{R}}$ and $L^1_{\mathbb{C}}$ to distinguish between the real and complex vector space versions of $L^1(\mu)$.

Theorem 1.24 (L^1 as a $\mathbb{C}$ normed vector space)**.** *The set $L^1_{\mathbb{C}}$ of is a vector space over $\mathbb{C}$ (with the usual addition and scalar multiplication of functions). Morever, if $f,g\in L^1_{\mathbb{C}}$ and $c\in\mathbb{C}$, then*

(a) the mapping $\Lambda : L^1\to\mathbb{C}$ defined by $\Lambda(f)=\int f$ is linear;
(b) $\left|\int f\right| \le \int |f|$.
(c) $\|cf\|_1 = |c|\|f\|_1$.
(d) $\|f+g\|_1 \le \|f\|_1 + \|g\|_1$.

Sketch of proof. Write $f=u+iv$ and $g=x+iy$. In particular, u,v,x,y are all $L^1_{\mathbb{R}}$. Given $c=a+ib$, the functions au, bv, av, bu are all $L^1_{\mathbb{R}}$ and so are $au-bv$ and $av+bu$ since $L^1_{\mathbb{R}}$ is a real vector space. Therefore, $cf=(au-bv)+i(av+bu)$ is in $L^1_{\mathbb{C}}$. A similar, but easier, argument shows $f+g$ is in $L^1_{\mathbb{C}}$. Hence $L^1_{\mathbb{C}}$ is a vector space over $\mathbb{C}$. Moreover, since the integral is real linear on $L^1_{\mathbb{R}}$,

$$\begin{aligned}\Lambda(cf) &= \Lambda((au+bv)+i(av+bu)) = \Lambda((au+bv)) + i\Lambda((av+bu))\\ &= a\Lambda(u)+b\Lambda(v)+i[a\Lambda(v)+b\Lambda(u)]\\ &= (a+ib)[\Lambda(u)+i\Lambda(v)] = c\Lambda(f).\end{aligned}$$

Likewise $\Lambda(f+g) = \Lambda(f) + \Lambda(g)$. Thus Λ is $\mathbb{C}$-linear on $L^1_{\mathbb{C}}$ and item (a) is proved.

If $\int f = 0$, then $f = 0$ almost everywhere and the last three items hold. Otherwise, write $\int f = re^{it}$ in polar coordinates and observe

$$e^{-it} \int f \in \mathbb{R}_+.$$

Thus, from the definition and linearity of the integral

$$\mathbb{R}_+ \ni e^{-it} \int f = \int e^{-it} f = \int \operatorname{real} e^{-it} f \, + \, i \int \operatorname{image} e^{-it} f.$$

Thus $\int \operatorname{image} e^{-it} f = 0$ and using results for $L^1_{\mathbb{R}}$,

$$\left| \int f \right| = \left| e^{-it} \int f \right| = \int e^{-it} f = \int \operatorname{real} e^{-it} f \leq \int |\operatorname{real} e^{-it} f| \leq \int |f|,$$

proving item (b).

Next,

$$\int |cf| = \int |c| \, |f| = |c| \int |f|.$$

Hence item (c) holds. Similarly,the triangle inequality, item (d), follows from $|f+g| \leq |f| + |g|$ (pointwise). $\square$

Remark 1.25. Proposition 1.24 says $\|\cdot\|_1$ is a semi-norm on L^1. As usual, we identify functions that differ by a null vector; that is, $f \sim g$ if $\|f - g\|_1 = 0$; equivalently, identifying functions that are equal a.e., we obtain a normed complex vector space of L^1 *functions* (which of course are not actually functions).

1.2.5. *Sequence spaces.* Define

$$\begin{aligned} c_0 &:= \{f : \mathbb{N} \to \mathbb{F} | \lim_{m \to \infty} |f(m)| = 0\} \\ \ell^\infty &:= \{f : \mathbb{N} \to \mathbb{F} | \sup_{m \in \mathbb{N}} |f(m)| < \infty\} \\ \ell^1 &:= \{f : \mathbb{N} \to \mathbb{F} | \sum_{m=0}^{\infty} |f(m)| < \infty\}. \end{aligned}$$

Note that $\ell^\infty = F_b(\mathbb{N}, \mathbb{F})$ and is a Banach space with the norm

$$\|f\|_\infty = \sup_m |f(m)|.$$

Further, $c_0 \subseteq \ell^\infty$ is the subspace $C_0(\mathbb{N})$ of ℓ^∞ again with the norm $\|\cdot\|_\infty$. In particular, c_0 is a Banach space.

Observe that ℓ^1 is the space $(\mathbb{N}, P(\mathbb{N}), c)$, where c is counting measure on $\mathbb{N}$ and $\|\cdot\|_1$ is the corresponding ℓ^1 norm. Since only set of measure zero in this measure space is the emptyset, two functions in $\ell^1 = L^1(c)$ are equivalent if and only if they are equal.

Along with these spaces it is also helpful to consider the vector space

$$c_{00} := \{f : \mathbb{N} \to \mathbb{F} | f(n) = 0 \text{ for all but finitely many } n\}$$

Notice that c_{00} is a vector subspace of each of c_0, ℓ^1 and ℓ^∞. Thus it can be equipped with either the $\|\cdot\|_\infty$ or $\|\cdot\|_1$ norms. It is not complete in either of these norms, however. What is true is that c_{00} is *dense* in c_0 and ℓ^1 (but not in ℓ^∞). (See Problem 1.11).

1.2.6. L^p *spaces.* Again let $(X, \mathscr{M}, m)$ be a measure space. For $1 \leq p < \infty$ let $L^p(m)$ denote the set of measurable functions f for which

$$\|f\|_p := \left(\int_X |f|^p \, dm\right)^{1/p} < \infty$$

(again we identify f and g when $f = g$ a.e.). It turns out that this quantity is a norm on $L^p(m)$, and $L^p(m)$ is complete, though we will not prove this yet (it is not immediately obvious that the triangle inequality holds when $p > 1$).

Choosing $(X, \mathscr{M}, \mu) = (\mathbb{N}, P(\mathbb{N}), c)$, counting measure on $\mathbb{N}$, obtains the sequence spaces ℓ^p; that is, the $\mathbb{F}$-vector space of functions $f : \mathbb{N} \to \mathbb{F}$ such that

$$\|f\|_p := \left(\sum_{n=1}^\infty |f(n)|^p\right)^{1/p} < \infty$$

and this quantity is a norm making ℓ^p into a Banach space.

When $p = \infty$, we define $L^\infty(\mu)$ to be the set of all functions $f : X \to \mathbb{F}$ with the following property: there exists $M > 0$ such that

$$(1) \qquad |f(x)| \leq M \quad \text{for } \mu-\text{a.e. } x \in X;$$

as for the other L^p spaces we identify f and g when there are equal a.e. When $f \in L^\infty$, let $\|f\|_\infty$ be the smallest M for which (1) holds. Then $\|\cdot\|_\infty$ is a norm making $L^\infty(\mu)$ into a Banach space.

1.2.7. *Subspaces and products.* If $(\mathcal{X}, \|\cdot\|)$ is a normed vector space and $\mathcal{Y} \subseteq \mathcal{X}$ is a vector subspace, then the restriction of $\|\cdot\|$ to $\mathcal{Y}$ is clearly a norm on $\mathcal{Y}$. If $\mathcal{X}$ is a Banach space, then $(\mathcal{Y}, \|\cdot\|)$ is a Banach space if and only if $\mathcal{Y}$ is *closed* in the norm topology of $\mathcal{X}$. (This is just a standard fact about metric spaces—a subspace of a complete metric space is complete in the restricted metric if and only if it is closed.)

Definition 1.26. A *subspace* $\mathcal{Y}$ of a normed vector space $\mathcal{X}$ is a closed vector subspace of $\mathcal{X}$, denoted $\mathcal{Y} \leq \mathcal{X}$. The terminology *linear manifold* is used synonymously with vector subspace.

If $\mathcal{X}, \mathcal{Y}$ are vector spaces then the *algebraic direct sum* is the vector space of ordered pairs

$$\mathcal{X} \oplus \mathcal{Y} := \{(x, y) : x \in \mathcal{X}, y \in \mathcal{Y}\}$$

with entrywise operations. If $\mathcal{X}$, $\mathcal{Y}$ are equipped with norms $\|\cdot\|_{\mathcal{X}}, \|\cdot\|_{\mathcal{Y}}$, then each of the quantities

$$\begin{aligned} \|(x,y)\|_\infty &:= \max(\|x\|_{\mathcal{X}}, \|y\|_{\mathcal{Y}}), \\ \|(x,y)\|_1 &:= \|x\|_{\mathcal{X}} + \|y\|_{\mathcal{Y}} \\ \|(x,y)\|_2 &:= \sqrt{\|x\|_{\mathcal{X}}^2 + \|y\|_{\mathcal{Y}}^2} \end{aligned} \tag{2}$$

is a norm on $\mathcal{X} \oplus \mathcal{Y}$. These three norms are equivalent; indeed it follows from the definitions that

$$\|(x,y)\|_\infty \leq \|(x,y)\|_2 \leq \|(x,y)\|_1 \leq 2\|(x,y)\|_\infty.$$

If $\mathcal{X}$ and $\mathcal{Y}$ are both complete, then $\mathcal{X} \oplus \mathcal{Y}$ is complete in each of these norms. The resulting Banach spaces are denoted $\mathcal{X} \oplus_\infty \mathcal{Y}$, $\mathcal{X} \oplus_1 \mathcal{Y}$ and $\mathcal{X} \oplus_2 \mathcal{Y}$.

Since the three norms in the previous paragraph are equivalent, the resulting spaces are indistinguishable topologically. There is a more abstract description of this topology.

Definition 1.27. *Given topological spaces (X, τ) and (Y, σ), the product topology on the Cartesian product $X \times Y$ is the smallest topoology that makes the coordinate projections $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ defined by $\pi_{\mathcal{X}}(x, y) = x, \pi_{\mathcal{Y}}(x, y) = y$ continuous. That is, the topology generated by the sets $U \times Y$ and $X \times V$ for open sets $U \subseteq X$ and $V \subseteq Y$.*

Proposition 1.28. *Suppose (X, τ) and (Y, σ) are topological spaces. The collection of sets*

$$\mathcal{B} = \{U \times V : U \subseteq X,\ V \subseteq Y \text{ are open}\}$$

is a base for the product topology.

For normed vector spaces $\mathcal{X}$ and $\mathcal{Y}$, the product topology on $\mathcal{X} \times \mathcal{Y}$ is metrizable and is the norm topology on $\mathcal{X} \times \mathcal{Y}$ with any of the norms of equation (2). Consequentely, a sequence $z_n = (x_n, y_n)$ from $\mathcal{X} \times \mathcal{Y}$ converges (in the product topology) if and only if both (x_n) and (y_n) converge; and z_n converges to $z = (x, y)$ if and only if (x_n) converges to x and (y_n) converges to y. In particular, if $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, then so is $\mathcal{X} \times \mathcal{Y}$.

It is evident how to extend the discussion here to finite products. The product topology is the default topology on (finite) products of Banach spaces (and more generally normed vector spaces).

1.2.8. *Qoutient spaces.* If $\mathcal{X}$ is a normed vector space and $\mathcal{M}$ is a proper subspace, then one can form the *algebraic quotient* $\mathcal{X}/\mathcal{M}$, defined as the collection of distinct cosets $\{x + \mathcal{M} : x \in \mathcal{X}\}$. From linear algebra, $\mathcal{X}/\mathcal{M}$ is a vector space under the standard operations. Let $\pi : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ denote the quotient map.

Proposition 1.29. *If $\mathcal{M}$ is a* closed *subspace of a normed vector space $\mathcal{X}$, then the quantity*

$$\|\pi(x)\| = \|x + \mathcal{M}\| := \inf_{y \in \mathcal{M}} \|x - y\|$$

is a norm on $\mathcal{X}/\mathcal{M}$. Moreover, if $\mathcal{X}$ is a Banach space, then so is $\mathcal{X}/\mathcal{M}$.

The norm in Proposition 1.29 is called the *quotient norm*. Geometrically, $\|x + \mathcal{M}\|$ is the *distance* in $\mathcal{X}$ from x to the closed set $\mathcal{M}$. The assumption that $\mathcal{M}$ is closed in needed to ensure that the quotient norm is indeed a norm. For instance $M = C([0,1])$ is dense subspace of $L^1([0,1])$ (with Lebesgue measure) and hence for any

$$\inf_{g \in \mathcal{M}} \|f - g\| = 0$$

for all $f \in L^1([0,1])$.

Proof. We will verify a couple of the axioms of a norm for the quotient norm, leaving the remainder of the proof as an exercise. First suppose $x \in \mathcal{X}$ and $\|\pi(x)\| = 0$. It follows that there is a sequence (m_n) form $\mathcal{M}$ such that $(\|x - m_n\|)$ converges to 0; that is, (m_n) converges to x. Since $\mathcal{M}$ is closed, $x \in \mathcal{M}$ and hence $\pi(x) = 0$.

Now let $x, y \in \mathcal{X}$ and $\epsilon > 0$ be given. There exists $m, n \in \mathcal{M}$ such that

$$\|x - m\| \leq \|\pi(x)\| + \epsilon, \quad \|y - n\| \leq \|\pi(y)\| + \epsilon.$$

Hence

$$\|\pi(x) + \pi(y)\| = \|\pi(x+y)\| \leq \|x + y - (m+n)\| \leq \|x - m\| + \|y - n\| \leq \|\pi(x)\| + \|\pi(y)\|| + 2\epsilon,$$

from which it follows that the triangle inequality holds and we have proved the quotient norm is indeed a norm.

To prove $\mathcal{X}/\mathcal{M}$ is complete (with the quotient norm) under the assumption that $\mathcal{X}$ is a Banach space (complete), suppose (y_n) is a sequence from $\mathcal{X}/\mathcal{M}$ and $\sum y_n$ is absolutely convergent. For each n there exists $x_n \in \mathcal{X}$ such that $\|x_n\| \leq \|y_n\| + \frac{1}{n^2}$ and $\pi(x_n) = y_n$. It follows that $\sum x_n$ is absolutely convergent. Since $\mathcal{X}$ is a Banach space the sequence of partial sums $s_N = \sum_{n=1}^{N} x_n$ converges to some $x \in \mathcal{X}$. In partiulcar,

$$\|s_N - x\| \geq \|\pi(s_N - x)\| = \|\pi(s_N) - \pi(x)\| = \|\sum_{n=1}^{N} y_n - \pi(x)\|.$$

Since $(\|s_N - x\|)$ converges to 0, it follows that $\sum y_n$ converges to $\pi(x)$. Hence $\mathcal{X}/\mathcal{M}$ is complete by Proposition 1.7. □

More examples are given in the exercises and further examples will appear after the development of some theory.

1.3. Linear transformations between normed spaces.

Definition 1.30. Let $\mathcal{X}, \mathcal{Y}$ be normed vector spaces. A linear transformation $T : \mathcal{X} \to \mathcal{Y}$ is *bounded* if there exists a constant $C \geq 0$ such that $\|Tx\|_{\mathcal{Y}} \leq C\|x\|_{\mathcal{X}}$ for all $x \in \mathcal{X}$.

Remark 1.31. Note that in Definition 1.30 it suffices to require that $\|Tx\|_{\mathcal{Y}} \leq C\|x\|_{\mathcal{X}}$ just for all $x \neq 0$, or for all x with $\|x\|_{\mathcal{X}} = 1$ (why?). □

The importance of boundedness and the following simple proposition is hard to overstate. Recall, a mapping $f : X \to Y$ between metric spaces is *Lipschitz continuous* if there is a constant $C > 0$ such that $d(f(x), f(y)) \leq Cd(x, y)$ for all $x, y \in X$. A simple exercise shows Lipschitz continuity implies (uniform) continuity.

Proposition 1.32. *If $T : \mathcal{X} \to \mathcal{Y}$ is a linear transformation between normed spaces, then the following are equivalent:*

(i) T is bounded.
(ii) T is Lipschitz continuous.
(iii) T is uniformly continuous.
(iv) T is continuous.
(v) T is continuous at 0.

Moreover, in this case,

$$\begin{aligned}\|T\| &:= \sup\{\|Tx\| : \|x\| = 1\}\\ &= \sup\{\frac{\|Tx\|}{\|x\|} : x \neq 0\}\\ &= \inf\{C : \|Tx\| \leq C\|x\| \text{ for all } x \in \mathcal{X}\}\end{aligned}$$

and $\|T\|$ is the smallest number (the infimum is attained [1]) such that

$$\|Tx\| \leq \|T\|\,\|x\| \tag{3}$$

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

Proof. Suppose T is bounded; that is, there exists a $C \geq 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in \mathcal{X}$. Thus, if $x, y \in \mathcal{X}$, then, $\|Tx - Ty\| = \|T(x - y)\| \leq C\|x - y\|$ by linearity of T. Hence (i) implies (ii). The implications (ii) implies (iii) implies (iv) implies (v) are evident.

The proof of (v) implies (i) exploits the homogeneity of the norm and the linearity of T and not nearly the full strength of the continuity assumption. By hypothesis, with $\epsilon = 1$ there exists $\delta > 0$ such that if $\|x\| = \|x - 0\| < 2\delta$, then $\|Tx\| = \|Tx - T0)\| < 1$. Given a nonzero vector $x \in \mathcal{X}$, the vector $\delta x/\|x\|$ has norm less than δ, so

$$1 > \left\|T\left(\frac{\delta x}{\|x\|}\right)\right\| = \delta\frac{\|Tx\|}{\|x\|}.$$

[1]The suprema need not be attained.

Rearranging this we find $\|Tx\| \leq (1/\delta)\|x\|$ for all $x \neq 0$.

Assuming T is bounded, it is immediate that $\sup\{\|Tx\| : \|x\| = 1\}$ exists (and is a real number). From homogeneity of the norm, it is also clear that

$$\sup\{\|Tx\| : \|x\| = 1\} = \sup\{\frac{\|Tx\|}{\|x\|} : x \neq 0\}.$$

Likewise assuming T is bounded the set $S = \{C : \|Tx\| \leq C\|x\| \text{ for all } x \in \mathcal{X}\} \subseteq [0, \infty)$ is not empty and bounded below (by 0) and hence the infimum exists. From the definition of $\|T\|$ we see that $\|T\| \in S$. Hence the infimum is at most $\|T\|$. On the other hand, if $C' < \inf S$, then there is an $x \in \mathcal{X}$ such that $\|Tx\| > C'\|x\|$ so that $\frac{\|Tx\|}{\|x\|} > C'$. Thus $C' < \|T\|$. ☐

The set of all bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ is denoted $B(\mathcal{X}, \mathcal{Y})$. It is a vector space under the operations of pointwise addition and scalar multiplication. The quantity $\|T\|$ is easily seen to be a norm. It is called the *operator norm* of T.

Problem 1.1. Prove the $\|\cdot\|_1$ and $\|\cdot\|_\infty$ norms on c_{00} are not equivalent. Conclude from your proof that the identity map on c_{00} is bounded from the $\|\cdot\|_1$ norm to the $\|\cdot\|_\infty$ norm, but not the other way around.

Problem 1.2. Consider c_0 and c_{00} equipped with the $\|\cdot\|_\infty$ norm. Prove there is no bounded operator $T : c_0 \to c_{00}$ such that $T|_{c_{00}}$ is the identity map. (Thus the conclusion of Proposition 1.34 can fail if $\mathcal{Y}$ is not complete.)

Proposition 1.33. *For normed vector spaces $\mathcal{X}$ and $\mathcal{Y}$, the operator norm makes $B(\mathcal{X}, \mathcal{Y})$ into a normed vector space that is complete if $\mathcal{Y}$ is complete.*

Proof. That $B(\mathcal{X}, \mathcal{Y})$ is a normed vector space follows readily from the definitions and is left as an exercise.

Suppose now $\mathcal{Y}$ is complete, and let T_n be a cauchy sequence in $B(\mathcal{X}, \mathcal{Y})$. Let $E = \mathcal{X}_1$ denote the closed unit ball in $\mathcal{X}$. For $x \in E$,

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\| \leq \|T_n - T_m\|. \tag{4}$$

Hence the sequence $(T_n|_E)$ is a cauchy sequence (so uniformly cauchy) from $C_b(E, \mathcal{Y})$, the space of bounded continuous functions from B to $\mathcal{Y}$. Since $\mathcal{Y}$ is complete, there is an $F \in C_b(E, \mathcal{Y})$ such that $(T_n|_E)$ converges to F in $C_b(E, \mathcal{Y})$ and moreover $\|F(x)\| \leq C := \sup\{\|T_n\| : n\} < \infty$. See Subsection 1.2.2. An exercise shows, given $x, y \in B$ and $c \in \mathbb{F}$ if $x + y \in B$ and $c\,x \in B$, then $F(x + y) = F(x) + F(y)$ and $F(cx) = cF(x)$. Hence F extends, by homogeneity, to a linear map $T : \mathcal{X} \to \mathcal{Y}$ such that $\|T\| \leq C$ and, by equation (4), (T_n) converges to T in $B(X, Y)$. ☐

If $T \in B(\mathcal{X}, \mathcal{Y})$ and $S \in B(\mathcal{Y}, \mathcal{Z})$, then two applications of the inequality (3) give, for $x \in \mathcal{X}$,

$$\|STx\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|$$

and it follows that $ST \in B(\mathcal{X}, \mathcal{Z})$ and $\|ST\| \leq \|S\|\|T\|$. In the special case that $\mathcal{Y} = \mathcal{X}$ is complete, $B(\mathcal{X}) := B(\mathcal{X}, \mathcal{X})$ is an example of a *Banach algebra*.

The following proposition is very useful in constructing bounded operators—at least when the codomain is complete. Namely, it suffices to define the operator (and show that it is bounded) on a dense subspace.

Proposition 1.34 (Extending bounded operators)**.** *Let $\mathcal{X}$, $\mathcal{Y}$ be normed vector spaces with $\mathcal{Y}$ complete, and $\mathcal{E} \subseteq \mathcal{X}$ a dense linear subspace. If $T : \mathcal{E} \to \mathcal{Y}$ is a bounded linear operator, then there exists a unique bounded linear operator $\widetilde{T} : \mathcal{X} \to \mathcal{Y}$ extending T (so $\widetilde{T}|_{\mathcal{E}} = T$). Further $\|\widetilde{T}\| = \|T\|$.*

Sketch of proof. Recall, if X, Y are metric spaces, Y is complete, $D \subseteq X$ is dense and $f : D \to Y$ is uniformly continuous, then f has a unique continuous extension $\tilde{f} : X \to Y$. Moreover, this extension can be defined as follows. Given $x \in X$, choose a sequence (x_n) from D converging to x and let $\tilde{f}(x) = \lim f(x_n)$ (that the sequence $f(x_n)$ is Cauchy follows from uniform continuity; that it converges from the assumption that $\mathcal{Y}$ is complete and finally it is an exercise to show $\tilde{f}(x)$ is well defined independent of the choice of (x_n)). Thus, it only remains to verify that the extension $\widetilde{T}$ of T is in fact linear and $\|T\| = \|\widetilde{T}\|$. Both are routine exercises. □

Example 1.35. Equip c_0 and c_{00} with the sup norm, $\|\cdot\|_\infty$ and consider the identity map $\iota : c_{00} \to c_{00}$. If T is an extension of ι to the completion c_0 of c_{00} (in the sup norm), then, letting $s_n \in c_{00}$ denote the sequence $s_n(m) = \frac{1}{m}$ for $m \leq n$ and $s_n(m) = 0$ for $m > n$, the sequence (s_n) is converge in c_0 to the sequence s with $s(m) = \frac{1}{m}$ for all m. Hence $(T(s_n) = s_n)$ converges to some $t \in c_{00}$. But now there is a K such that $t(k) = 0$ for all $k \geq K$ so that $\|s_n - t\| \geq \frac{1}{K}$ for all $n \geq K$, a contradiction. This example shows completeness of $\mathcal{Y}$ is essential in Proposition 1.34. □

Definition 1.36. A bounded linear transformation $T \in B(\mathcal{X}, \mathcal{Y})$ is said to be *invertible* if it is bijective (being bijective, automatically T^{-1} exists and is a linear transformation) and T^{-1} is bounded from $\mathcal{Y}$ to $\mathcal{X}$. Two normed spaces $\mathcal{X}, \mathcal{Y}$ are said to be *(boundedly) isomorphic* if there exists an invertible linear transformation $T : \mathcal{X} \to \mathcal{Y}$.

Example 1.37. As an example, given equivalent norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space $\mathcal{X}$, the identity mapping $\iota : (\mathcal{X}, \|\cdot\|_1) \to (\mathcal{X}, \|\cdot\|_2)$ is (boundedly) invertible and witnesses the fact that these two normed vector spaces are boundedly isomorphic.

Definition 1.38. An operator $T : \mathcal{X} \to \mathcal{Y}$ such that $\|Tx\| = \|x\|$ for all $x \in \mathcal{X}$ is an *isometry*. Note that an isometry is automatically injective and if it is also surjective then it is automatically invertible and T^{-1} is also an isometry. The normed vector spaces are *isometrically isomorphic* if there is an invertible isometry $T : \mathcal{X} \to \mathcal{Y}$.

Example 1.39. If $\mathcal{X}$ is a finite dimensional vector space and $T : \mathcal{X} \to \mathcal{X}$ is an isometry, then T is onto. However, when $\mathcal{X}$ is not finite dimensional, an isometry need not be surjective. As examples, let $\ell^p = \ell^p(\mathbb{N})$ denote the sequence spaces from Subsection 1.2.5. The linear map $S : \ell^p(\mathbb{N}) \to \ell^p(\mathbb{N})$ defined by $Sf(n) = 0$ if $n = 0$ and $f(n-1)$ if $n > 0$ (for $f = (f(n))_n \in \ell^p$) is the *shift operator*. It is straightforward to verify that S is an isometry but not onto.

Example 1.40. Following up on the previous example, a linear map $T : \mathcal{X} \to \mathcal{X}$ can be one-one and have dense range without being (boundedly) invertible. Let $e_n \in \ell^2(\mathbb{N})$ denote the function $e_n(m) = 1$ if $n = m$ and 0 otherwise for non-negative integers $0 \le m, n$. The set of $c_{00} = \{\sum_{n=0}^N a_n e_n : N \in \mathbb{N},\, c_n \in \mathbb{F}\}$ is dense in $\ell^2(\mathbb{N})$ and the mapping $D : c_{00} \to \ell^2(\mathbb{N})$ defined by

$$D(\sum_{n=0}^N a_n e_n) = \sum_{n=0}^N \frac{a_n}{n+1} e_n$$

is easily seen to be bounded with $\|D\| = 1$. It is also injective. Hence D extends to an injective bounded operator, still denoted D, from $\ell^2 \to \ell^2$, with $\|D\| = 1$. The range of D contains $\{e_n : n \in \mathbb{N}\}$ and is thus dense in $\ell^2(\mathbb{N})$.

Since $\sum_{n=0}^\infty |\frac{1}{n+1}|^2 < \infty$, the vector $f = \sum_{n=1}^\infty \frac{1}{n+1} e_n$ is in $\ell^2(\mathbb{N})$. On the other hand, if $g \in \ell^2(\mathbb{N})$ and $Dg = f$, then

$$\frac{1}{n+1} g(n) = (Dg)(n) = f(n) = \frac{1}{n+1}$$

and thus $g(n) = 1$ for all n; however, since $\sum |g(n)|^2 = \infty$, we obtain a contradiction. Hence f is not in the range of D.

1.4. **Examples.**

(a) If $\mathcal{X}$ is a finite-dimensional normed space and $\mathcal{Y}$ is any normed space, then every linear transformation $T : \mathcal{X} \to \mathcal{Y}$ is bounded. See Problem 1.16.

(b) Let $\mathcal{X}$ denote c_{00} equipped with the $\|\cdot\|_1$ norm, and $\mathcal{Y}$ denote c_{00} equipped with the $\|\cdot\|_\infty$ norm. Then the identity map $id_{\mathcal{X},\mathcal{Y}} : \mathcal{X} \to \mathcal{Y}$ is bounded (in fact its norm is equal to 1), but its inverse, the identity map $\iota_{\mathcal{Y},\mathcal{X}} : \mathcal{Y} \to \mathcal{X}$, is unbounded. To verify this claim, For positive integers n, let f_n denote the element of c_{00} defined by

$$f_n(m) = \begin{cases} 1 & \text{if } m \le n \\ 0 & \text{if } m > n. \end{cases}$$

Now $\|\iota_{\mathcal{Y},\mathcal{X}}(f_n)\|_1 = n$, but $\|f_n\|_\infty = 1$.

(c) Consider c_{00} with the $\|\cdot\|_\infty$ norm. Let $a : \mathbb{N} \to \mathbb{F}$ be any function and define a linear transformation $T_a : c_{00} \to c_{00}$ by

$$T_a f(n) = a(n) f(n). \tag{5}$$

The mapping T_a is bounded if and only if $M = \sup_{n\in\mathbb{N}} |a(n)| < \infty$, in which case $\|T_a\| = M$. In this case, T_a extends uniquely to a bounded operator from c_0 to c_0

by Proposition 1.34, and one may check that the formula (5) defines the extension. All of these claims remain true if we use the $\|\cdot\|_1$ norm instead of the $\|\cdot\|_\infty$ norm. In this case, we get a bounded operator from ℓ^1 to itself.

(d) Define $S : \ell^1 \to \ell^1$ as follows: given the sequence $(f(n))_n$ from ℓ^1 let $Sf(1) = 0$ and $Sf(n) = f(n-1)$ for $n > 1$. (Viewing f as a sequence, S shifts the sequence one place to the right and fills in a 0 in the first position). This S is an isometry, but is not surjective. In contrast, if $\mathcal{X}$ is finite-dimensional, then the rank-nullity theorem from linear algebra guarantees that every injective linear map $T : \mathcal{X} \to \mathcal{X}$ is also surjective.

(e) Let $C^\infty([0,1])$ denote the vector space of functions on $[0,1]$ with continuous derivatives of all orders. The differentiation map $D : C^\infty([0,1]) \to C^\infty([0,1])$ defined by $Df = \frac{df}{dx}$ is a linear transformation. Since, for $t \in \mathbb{R}$, we have $De^{tx} = te^{tx}$, it follows that there is *no* norm on $C^\infty([0,1])$ such that $\frac{d}{dx}$ is bounded.

1.5. Problems.

Problem 1.3. Prove Proposition 1.5.

Problem 1.4. Prove equivalent norms define the same topology and the same Cauchy sequences.

Problem 1.5. (a) Prove all norms on a finite dimensional vector space $\mathcal{X}$ are equivalent. Suggestion: Fix a basis $e_1, \dots e_n$ for $\mathcal{X}$ and define $\|\sum a_k e_k\|_1 := \sum |a_k|$. It is routine to check that $\|\cdot\|_1$ is a norm on $\mathcal{X}$. Now complete the following outline.

- (i) Let $\|\cdot\|$ be the given norm on $\mathcal{X}$. Show there is an M such that $\|x\| \le M\|x\|_1$. Conclude that the mapping $\iota : (\mathcal{X}, \|\cdot\|_1) \to (\mathcal{X}, \|\cdot\|)$ defined by $\iota(x) = x$ is continuous;
- (ii) Show that the *unit sphere* $S = \{x \in \mathcal{X} : \|x\|_1 = 1\}$ in $(\mathcal{X}, \|\cdot\|_1)$ is compact in the $\|\cdot\|_1$ topology;
- (iii) Show that the mapping $f : S \to (\mathcal{X}, \|\cdot\|)$ given by $f(x) = \|x\|$ is continuous and hence attains its infimum. Show this infimum is not 0 and finish the proof.

(b) Combine the result of part (a) with the result of Problem 1.4 to conclude that every finite-dimensional normed vector space is complete.

(c) Let $\mathcal{X}$ be a normed vector space and $\mathcal{M} \subseteq \mathcal{X}$ a finite-dimensional subspace. Prove $\mathcal{M}$ is closed in $\mathcal{X}$.

Problem 1.6. Finish the proofs from the examples subsections.

Problem 1.7. A function $f : [0,1] \to \mathbb{F}$ is called *Lipschitz continuous* if there exists a constant C such that

$$|f(x) - f(y)| \le C|x - y|$$

for all $x, y \in [0,1]$. Define $\|f\|_{Lip}$ to be the best possible constant in this inequality. That is,

$$\|f\|_{Lip} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

Let $\mathrm{Lip}[0,1]$ denote the set of all Lipschitz continuous functions on $[0,1]$. Prove $\|f\| := |f(0)| + \|f\|_{Lip}$ is a norm on $\mathrm{Lip}[0,1]$, and that $\mathrm{Lip}[0,1]$ is complete in this norm.

Problem 1.8. Let $C^1([0,1])$ denote the space of all functions $f : [0,1] \to \mathbb{R}$ such that f is differentiable in $(0,1)$ and f' extends continuously to $[0,1]$. Prove

$$\|f\| := \|f\|_\infty + \|f'\|_\infty$$

is a norm on $C^1([0,1])$ and that C^1 is complete in this norm. Do the same for the norm $\|f\| := |f(0)| + \|f'\|_\infty$. (Is $\|f'\|_\infty$ a norm on C^1?)

Problem 1.9. Let $(X, \mathscr{M})$ be a measurable space. Let $M(X)$ denote the (real) vector space of all signed measures on $(X, \mathscr{M})$. Prove the *total variation norm* $\|\mu\| := |\mu|(X)$ is a norm on $M(X)$, and $M(X)$ is complete in this norm.

Problem 1.10. Prove, if $\mathcal{X}, \mathcal{Y}$ are normed spaces, then the operator norm is a norm on $B(\mathcal{X}, \mathcal{Y})$.

Problem 1.11. Prove c_{00} is dense in c_0 and ℓ^1. (That is, given $f \in c_0$ there is a sequence f_n in c_{00} such that $\|f_n - f\|_\infty \to 0$, and the analogous statement for ℓ^1.) Using these facts, or otherwise, prove that c_{00} is *not* dense in ℓ^∞. (In fact there exists $f \in \ell^\infty$ with $\|f\|_\infty = 1$ such that $\|f - g\|_\infty \geq 1$ for all $g \in c_{00}$.)

Problem 1.12. Prove c_{00} is not complete in the $\|\cdot\|_1$ or $\|\cdot\|_\infty$ norms. (After we have studied the Baire Category theorem, you will be asked to prove that there is *no* norm on c_{00} making it complete.)

Problem 1.13. Consider c_0 and c_{00} equipped with the $\|\cdot\|_\infty$ norm. Prove there is no bounded operator $T : c_0 \to c_{00}$ such that $T|_{c_{00}}$ is the identity map. (Thus the conclusion of Proposition 1.34 can fail if $\mathcal{Y}$ is not complete.)

Problem 1.14. Prove the $\|\cdot\|_1$ and $\|\cdot\|_\infty$ norms on c_{00} are not equivalent. Conclude from your proof that the identity map on c_{00} is bounded from the $\|\cdot\|_1$ norm to the $\|\cdot\|_\infty$ norm, but not the other way around.

Problem 1.15. a) Prove $f \in C_0(\mathbb{R}^n)$ if and only if f is continuous and $\lim_{|x|\to\infty} |f(x)| = 0$. b) Let $C_c(\mathbb{R}^n)$ denote the set of continuous, compactly supported functions on $\mathbb{R}^n$. Prove $C_c(\mathbb{R}^n)$ is dense in $C_0(\mathbb{R}^n)$ (where $C_0(\mathbb{R}^n)$ is equipped with sup norm).

Problem 1.16. Prove, if $\mathcal{X}, \mathcal{Y}$ are normed spaces and $\mathcal{X}$ is finite dimensional, then every linear transformation $T : \mathcal{X} \to \mathcal{Y}$ is bounded. Suggestion: Let d denote the dimension of X and let $\{e_1, \ldots, e_d\}$ denote a basis. The function $\|\cdot\|_1$ on $\mathcal{X}$ defined by $\|\sum x_j e_j\|_1 = \sum |x_j|$ is a norm. Apply Problem 1.5.

Problem 1.17. Prove the claims in Example 1.4(c).

Problem 1.18. Let $g : \mathbb{R} \to \mathbb{F}$ be a (Lebesgue) measurable function. The map $Mg : f \to gf$ is a linear transformation on the space of measurable functions. Prove, if $g \notin L^\infty(\mathbb{R})$, then there is an $f \in L^1(\mathbb{R})$ such that $gf \notin L^1(\mathbb{R})$. Conversely, show if $g \in L^\infty(\mathbb{R})$, then M_g is bounded from $L^1(\mathbb{R})$ to itself and $\|M_g\| = \|g\|_\infty$.

Problem 1.19. Prove the claims about direct sums.

Problem 1.20. Let $\mathcal{X}$ be a normed vector space and $\mathcal{M}$ a proper *closed* subspace. Prove for every $\epsilon > 0$, there exists $x \in \mathcal{X}$ such that $\|x\| = 1$ and $\inf_{y\in\mathcal{M}} \|x - y\| > 1 - \epsilon$. (Hint: take any $u \in \mathcal{X} \setminus \mathcal{M}$ and let $a = \inf_{y\in\mathcal{M}} \|u - y\|$. Choose $\delta > 0$ small enough so that $\frac{a}{a+\delta} > 1-\epsilon$, and then choose $v \in \mathcal{M}$ so that $\|u-v\| < a+\delta$. Finally let $x = \frac{u-v}{\|u-v\|}$.)

Note that the distance to a (closed) subspace need not be attained. Here is an example. Consider the Banach space $C([0, 1])$ (with the sup norm of course and either real or complex valued functions) and the closed subspace

$$T = \{f \in C([0, 1]) : f(0) = 0 = \int_0^1 f\, dt\}.$$

Using machinery in the next section it will be evident that T is a closed subspace of $C([0, 1])$. For now, it can be easily verified directly. Let g denote the function $g(t) = t$. Verify that, for $f \in T$, that

$$\frac{1}{2} = \int g\, dt = \int (g - f)\, dt \leq \|g - f\|_\infty.$$

In particular, the distance from g to T is at least $\frac{1}{2}$.

Note that the function $h = x - \frac{1}{2}$, while not in T, satisfies $\|g - h\|_\infty = \frac{1}{2}$.

On the other hand, for any $\epsilon > 0$ there is an $f \in T$ so that $\|g - f\|_\infty \leq \frac{1}{2} + \epsilon$ (simply modify h appropriately). Thus, the distance from g to T is $\frac{1}{2}$. Now verify, using the inequality above, that h is the only element of $C([0, 1])$ such that $\int h\, dt = 0$ and $\|g - h\|_\infty = \frac{1}{2}$.

Problem 1.21. Prove, if $\mathcal{X}$ is an infinite-dimensional normed space, then the unit ball $ball(\mathcal{X}) := \{x \in \mathcal{X} : \|x\| \leq 1\}$ is not compact in the norm topology. (Hint: use the result of Problem 1.20 to construct inductively a sequence of vectors $x_n \in \mathcal{X}$ such that $\|x_n\| = 1$ for all n and $\|x_n - x_m\| \geq \frac{1}{2}$ for all $m < n$.)

Problem 1.22. (The quotient norm) Let $\mathcal{X}$ be a normed space and $\mathcal{M}$ a proper closed subspace.

a) Prove the quotient norm is a norm.
b) Show that the quotient map $x \to x + \mathcal{M}$ has norm 1. (Use Problem 1.20.)
c) Prove, if $\mathcal{X}$ is complete, so is $\mathcal{X}/\mathcal{M}$.

Problem 1.23. A normed vector space $\mathcal{X}$ is called *separable* if it is separable as a metric space (that is, there is a countable subset of $\mathcal{X}$ which is dense in the norm topology). Prove c_0 and ℓ^1 are separable, but ℓ^∞ is not. (Hint: for ℓ^∞, show that there is an uncountable collection of elements $\{f_\alpha\}$ such that $\|f_\alpha - f_\beta\| = 1$ for $\alpha \neq \beta$.)

END FALL TERM

2. Linear functionals and the Hahn-Banach theorem

If there is a *fundamental theorem of functional analysis,* it is the Hahn-Banach theorem. The theorem is somewhat abstract-looking at first, but its importance will be clear after studying some of its corollaries.

Definition 2.1. Let $\mathcal{X}$ be a normed vector space over the field $\mathbb{F}$. A *linear functional* on $\mathcal{X}$ is a linear map $L : \mathcal{X} \to \mathbb{F}$. The *dual space* of $\mathcal{X}$, denoted $\mathcal{X}^*$ is the space $B(\mathcal{X}, \mathbb{F})$ of bounded linear functionals on $\mathcal{X}$.

Remark 2.2. Since $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is complete, the vector space of bounded linear functionals is itself a Banach space (complete normed vector space) and is known as the . It is not yet obvious that $\mathcal{X}^*$ need be non-trivial (that is, that there are any bounded linear functionals on $\mathcal{X}$ besides 0). One corollary of the Hahn-Banach theorem is there exist enough bounded linear functionals on $\mathcal{X}$ to separate points.

2.1. Examples.

This subsection contains some examples of bounded linear functionals and dual spaces.

Example 2.3. For each of the sequence spaces c_0, ℓ^1, ℓ^∞, for each n the map $f \to f(n)$ is a bounded linear functional. That is, $\lambda_n : \mathcal{X} \to \mathbb{F}$ defined by $\lambda_n(f) = f(n)$ for $f : \mathbb{N} \to \mathbb{F}$ in $\mathcal{X}$, where $\mathcal{X}$ is any one of c_0, ℓ^1, ℓ^∞, is continuous since in each case it is immediate that

$$|\lambda_n(f)| = |f(n)| \leq \|f\|_{\mathcal{X}}.$$

Example 2.4. *Given $g \in \ell^1$, if $f \in c_0$, then*

$$\sum_{n=0}^{\infty} |f(n)g(n)| \leq \|f\|_\infty \sum_{n=0}^{\infty} |g(n)| = \|g\|_1 \|f\|_\infty. \tag{6}$$

Thus $\sum_{n=0}^{\infty} f(n)g(n)$ converges and we obtain a functional $L_g : c_0 \to \mathbb{F}$ defined by

$$L_g(f) := \sum_{n=0}^{\infty} f(n)g(n). \tag{7}$$

The inequality of equation (6) says L_g is bounded (continuous) and $\|L_g\| \leq \|g\|_1$. Moreover, it is immediate that $\Phi : \ell^1 \to c_0^*$ defined by $\Phi(g) = L_g$ is bounded and linear and $\|\Phi\| \leq 1$. In fact, Φ is onto so that every bounded linear functional on c_0 is of the form L_g for some $g \in \ell^1$.

Proposition 2.5. *The map $\Phi : \ell^1 \to c_0^*$ defined by $\Phi(g) = L_g$ is an isometric isomorphism from ℓ^1 onto the dual space c_0^*.*

Proof. We have already seen that each $g \in \ell^1$ gives rise to a bounded linear functional $L_g \in c_0^*$ via

$$L_g(f) := \sum_{n=0}^{\infty} g(n)f(n),$$

that $\|L_g\| \leq \|g\|_1$ and the the mapping Φ is bounded and linear. We will prove simultaneously that this map is onto and that $\|L_g\| \geq \|g\|_1$.

Let $L \in c_0^*$. We will first show that there is unique $g \in \ell^1$ so that $L = L_g$. Let $e_n \in c_0$ be the indicator function of n, that is

$$e_n(m) = \delta_{nm}.$$

Define a function $g : \mathbb{N} \to \mathbb{F}$ by

$$g(n) = L(e_n).$$

We claim that $g \in \ell^1$ and $L = L_g$. To see this, fix an integer N and define $h = h_N : \mathbb{N} \to \mathbb{F}$ by

$$h(n) = \begin{cases} \overline{g(n)}/|g(n)| & \text{if } n \leq N \text{ and } g(n) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Thus $h = \sum_{n=0}^{N} h(n)e_n$. Further, by $h \in c_{00} \subseteq c_0$ and $\|h\|_\infty \leq 1$. Now

$$\sum_{n=0}^{N} |g(n)| = \sum_{n=0}^{N} h(n)g(n) = L(h) = |L(h)| \leq \|L\|\|h\| \leq \|L\|.$$

It follows that $g \in \ell^1$ and $\|g\|_1 \leq \|L\|$. By construction $L = L_g$ when restricted to c_{00}, so by the uniqueness of extensions of bounded operators, Proposition 1.34, $L = L_g$. Thus the map $g \to L_g$ is onto and

$$\|g\|_1 \leq \|L\| = \|L_g\| \leq \|g\|_1.$$

□

Example 2.6. Given $g \in \ell^\infty$, if $f \in \ell^1$, then equation (6) shows $|L_g(f)| \leq \|g\|_\infty \|f\|_1$, where L_g is defined as in equation (7). Thus $\|L_g\| \leq \|g\|_\infty$ and we obtain a bounded linear map $\Psi : \ell^\infty \to (\ell^1)^*$

Proposition 2.7. *The map Ψ is an isometric isomorphism from $(\ell^1)^*$ onto ℓ^∞.*

Proof. The proof follows the same lines as the proof of the previous proposition; the details are left as an exercise. □

Remark 2.8. The same mapping $g \to L_g$ also shows that every $g \in \ell^1$ gives a bounded linear functional on ℓ^∞, but it turns out these do not exhaust $(\ell^\infty)^*$ (see Problem 2.12).

Regarding ℓ^1 and ℓ^∞ as L^1 and L^∞ for counting measure on $\mathbb{N}$, it is not surprising that, given a measure space $(X, \mathscr{M}, \mu)$, a function $g \in L^\infty(\mu)$ (see Subsection 1.2.6 for the definition of $L^\infty(\mu)$) defines a linear functional $L_g : L^1(\mu) \to \mathbb{F}$ by

$$L_g(f) := \int_X fg\, dm$$

for $f \in L^1(\mu)$ is a bounded linear functional of norm at most M. We will prove in Section 4 that the norm of L_g is in fact $\|g\|_\infty$, and every bounded linear functional on $L^1(m)$ is of this type (at least when m is σ-finite). □

Example 2.9. A regular Borel measure μ on a locally compact set X such that $\mu(K) < \infty$ for compact subsets of X determines a linear functional $\lambda : C_c(X) \to \mathbb{F}$ by

$$\lambda(f) = \lambda_\mu(f) = \int_X f\, d\mu.$$

An $f \in C_c(X)$ is a *positive function* (really non-negative), written $f \geq 0$, if $f(x) \geq 0$ for all $x \in X$. The linear functional λ_μ is a *positive linear functional* in the sense that if $f \in C_c(X)$ is positive, then $\lambda_\mu(f) \geq 0$.

As a second example, let $X = [0,1]$ and note that the mapping $I : C([0,1]) \to \mathbb{C}$ defined by

$$I(f) = \int_0^1 f\, dx,$$

where the integral is in the Riemann sense, is a positive linear functional on $C([0,1])$.

END Monday 2025-01-13

Theorem 2.10 (Riesz-Markov Representation Theorem: positive version)**.** *Let $X = (X, \tau)$ be a locally compact Hausdorff space. If $\lambda : C_c(X) \to \mathbb{C}$ is a positive linear functional, then there exists a unique Borel measure μ on the Borel σ-algebra $\mathscr{B}_X$, such that*

$$\lambda(f) = \int f\, d\mu$$

for $f \in C_c(X)$. Moreover, μ is regular *in the sense that*

(i) if $K \subseteq X$ is compact, then $\mu(K) < \infty$;
(ii) if $E \in \mathscr{B}_X$, then $\mu(E) = \inf\{\mu(U) : E \subseteq U,\ U \text{ open}\}$; and
(iii) if $E \in \mathscr{B}_X$ and $\mu(E) < \infty$, then $\mu(E) = \sup\{\mu(K) : K \subseteq E,\ K \text{ compact}\}$.

Remark 2.11. In general elements of $C_c(X)^*$ correspond to *signed measures* that will appear later in these notes. □

2.2. **Continuous linear functionals.** For linear functionals, we can add to the list of equivalent conditions of Proposition 1.32. In particular, the proof of the equivalence of items (a) and (c) in Proposition 2.12 requires a map into the scalar field.

Proposition 2.12. *If $\mathcal{X}$ is a normed vector space and if $\lambda : \mathcal{X} \to \mathbb{F}$ is a non-zero (not identically zero) linear functional , then the following are equivalent.*

(a) λ *is continuous;*
(b) λ *is bounded;*
(c) $\ker \lambda$ *is closed;*
(d) $\overline{\ker \lambda} \neq \mathcal{X}$.

Proof. Items (a) and (b) are equivalent by Proposition 1.32 and it is evident that item (a) implies item (c).

Suppose item (b) does not hold. Thus, there exists a sequence (f_n) from $\mathcal{X}$ such that $\|f_n\| \leq 1$, but $|\lambda(f_n)| \geq n$. Choose $e \in \mathcal{X}$ with $\lambda(e) = 1$ and let

$$h_k = e - \frac{f_n}{\lambda(f_n)}$$

and note that (h_k) converges to e but $\lambda(h_k) = 0$ for all k. Thus (h_k) is a sequence from $\ker \lambda$ that converges to a point not in $\ker \lambda$. Thus item (c) implies item (b).

Since $\ker \lambda \neq V$ (since λ is not the zero map), item (c) implies item (d). Now suppose item (c) does not hold. Thus there exists an $f \notin \ker \lambda$ and a sequence (f_n) from $\ker \lambda$ that converges to f. Without loss of generality, $\lambda(f) = 1$. Given $g \in \mathcal{X}$, the sequence

$$g_n = (g - \lambda(g)f) + \lambda(g)\, f_n$$

converges to g and $\lambda(g_n) = 0$. Thus $g \in \overline{\ker \lambda}$ and we conclude $\mathcal{X} = \overline{\ker \lambda}$. ☐

Remark 2.13. Note that Proposition 2.12 remains true with $\lambda^{-1}(\{a\})$ in place of $\ker \lambda$, for any choice of $a \in \mathbb{F}$. For instance, in the proof that item (c) implies item (b), simply require $\lambda(e) = a + 1$ instead of $\lambda(e) = a/$ In the proof that item (d) implies item (c), suppose the sequence (f_n) converges to f and $\lambda(f_n) = a$, but $\lambda(f_n) = b \neq a$. In this case, given $g \in \mathcal{X}$, let $g_n = (g - cf) + cf_n <$ where $c = \frac{a-\lambda(g)}{a-b}$. The details are left as an exercise.

As a corollary, Proposition 2.12 extends to linear maps from a normed space $\mathcal{X}$ into a finite dimensional normed space as an easy argument shows.

If $\mathcal{X}$ is infinite dimensional, the result is false. Just choose a basis B for $\mathcal{X}$ and let $B_0 = \{b_1, b_2, \dots\}$ denote a countable subset of B. Define $L : \mathcal{X} \to \mathcal{X}$ by declaring $L(b_n) = n\, b_n$ and $L(b) = b$ for $b \in B \setminus B_0$ and extend by linearity. Thus L is one-one so that $\ker L = \{0\}$ is closed, but L is not bounded (and so not continuous). ☐

We close this subsection with the following result that should be compared with item (a) from Subsection 1.4 followed by an example.

Proposition 2.14. *If V is an infinite dimensional normed vector space, then there exists a linear map $f : V \to \mathbb{F}$ that is not continuous.*

For a Banach space $\mathcal{X}$, there are notions of a basis that reference the norm. For instance, a *Schauder basis* for $\mathcal{X}$ is a sequence $(e_n)_{n=1}^{\infty}$ such that for each $x \in \mathcal{X}$ there exists a unique choice of scalars $x_n \in \mathbb{F}$ such that the series $\sum_{n=1}^{\infty} x_n e_n$ converges to x. Forgetting the norm structure, a *Hamel basis* $B \subseteq \mathcal{X}$ for $\mathcal{X}$ is a basis in the sense of linear algebra. Explicitly, letting $F_{00}(B)$ denote the functions $a : B \to \mathbb{F}$ such that $a_b = a(b)$ is zero for all but finitely many $b \in B$, the set B is a Hamel basis for $\mathcal{X}$ if for each $v \in \mathcal{X}$ there exist is a unique function $a \in F_{00}(B)$ such that

$$v = \sum_{b \in B} a_b\, b = \sum_{b \in B}^{\text{finite}} a_b\, b.$$

In this case any choice of $c : B \to \mathbb{F}$ determines uniquely a linear functional $\lambda : \mathcal{X} \to \mathbb{F}$ via the rule

$$\lambda(v) = \sum_{b \in B} c_b\, a_b,$$

where $c_b = c(b)$. Often this process is described informally as: *let* $\lambda(b) = c(b)$ *and extend by linearity.* Finally, an argument using Zorn's Lemma, which we will soon encounter in the proof of the Hahn-Banach Theorem, shows that every vector space has a basis. While it is true that every basis for a vector space V has the same cardinality, all that we need to make sense of the statement V *is an infinite dimensional vector space* is the fact that V has a basis that is infinite, then all bases for V are infinite, which is an immediate consequence of the fact that all bases for a finite dimensional vector space have the same cardinality. Thus, we can take the statement $\mathcal{X}$ *is infinite dimensional* to mean that $\mathcal{X}$ has a Hamel basis B that contains a countable set B_0.

Proof of Proposition 2.14. Let B denote a Hamel basis for V. By assumption, B has a countable subset B_0. Write $B_0 = \{b_1, b_2, \dots\}$ (so choose a bijection $\psi : \mathbb{N} \to B_0$) and assume, without loss of generality that $\|b_j\| = 1$. Let $\lambda : V \to \mathbb{F}$ denote the linear functional determined by $\lambda(b_j) = j$ for $b_j \in B_0$ and $\lambda(b) = 0$ for $b \in B \setminus B_0$ and observe that λ is not bounded. ☐

Example 2.15. Let $\mathcal{X}$ denote an infinite dimensional normed vector space and suppose $f : \mathcal{X} \to \mathbb{F}$ denote a discontinuous linear functional. An exercise shows that $\overline{f^{-1}(\{1\})} = \mathcal{X}$ in addition to $\overline{\ker f} = \mathcal{X}$ (see Proposition 2.12). Let $U = \{f \le 0\} \subseteq \ker \lambda$ and note $V = \mathcal{X} \setminus U = \{f > 0\} \supseteq f^{-1}(\{1\})$. Now U and V are disjoint convex sets such that $\mathcal{X} = U \cup V$ and $\overline{U} = \mathcal{X} = \overline{V}$.

2.3. **The Hahn-Banach Extension Theorem.** To state and prove the Hahn-Banach Extension Theorem, we first work in the setting $\mathbb{F} = \mathbb{R}$, then extend the results to the complex case.

Definition 2.16. Let $\mathcal{X}$ be a real vector space. A *Minkowski functional* is a function $p : \mathcal{X} \to \mathbb{R}$ such that $p(x+y) \le p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for all $x, y \in \mathcal{X}$ and nonnegative $\lambda \in \mathbb{R}$.

For examples, if $L : \mathcal{X} \to \mathbb{R}$ is any linear functional, then the function $p : \mathcal{X} \to \mathbb{R}$ defined by $p(x) := |L(x)|$ is a Minkowski functional; and if $\|\cdot\|$ is a seminorm on $\mathcal{X}$, then $p : \mathcal{X} \to \mathbb{R}$ defined by $p(x) = \|x\|$ is a Minkowski functional.

Theorem 2.17 (The Hahn-Banach Extension[2] Theorem, real version)**.** *Let $\mathcal{X}$ denote a real vector space, p a Minkowski functional on $\mathcal{X}$, and $\mathcal{M}$ a subspace of $\mathcal{X}$. If L is a linear functional on $\mathcal{M}$ such that $L(x) \le p(x)$ for all $x \in \mathcal{M}$, then there exists a linear functional L' on $\mathcal{X}$ such that*

(i) $L'|_{\mathcal{M}} = L$ (L' extends L)
(ii) $L'(x) \le p(x)$ for all $x \in \mathcal{X}$ (L' is dominated by p).

Remark 2.18. In the statement of Theorem 2.17, $\mathcal{X}$ is a vector space, not a normed vector space and correspondingly $\mathcal{M}$ is a subspace in the sense of linear algebra (sometimes referred to as a linear manifold). □

The proof will invoke *Zorn's Lemma*, a result that is equivalent to the axiom of choice (as well as the well-ordering principle and the Hausdorff maximality principle). A *partial order* $\preceq$ on a set S is a relation that is reflexive, symmetric and transitive; that is, for all $x, y, z \in S$

(i) $x \preceq x$,
(ii) if $x \preceq y$ and $y \preceq x$, then $x = y$, and
(iii) if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

We call $\mathcal{S}$, or more precisely $(S, \preceq)$, a *partially ordered set* or *poset*. A subset T of S is *totally ordered*, if for each $x, y \in T$ either $x \preceq y$ or $y \preceq x$. A totally ordered subset T is often called a *chain*. An upper bound z for a chain T is an element $z \in S$ such that $t \preceq z$ for all $t \in T$. A maximal element for S is a $w \in S$ that has no successor; that is there does not exist an $s \in S$ such that $s \ne w$ and $w \preceq s$. An *upper bound* for a subset A of S is an element $s \in S$ such that $a \preceq s$ for all $a \in A$.

Theorem 2.19 (Zorn's Lemma)**.** *Suppose S is a partially ordered set. If every chain in S has an upper bound, then S has a maximal element.*

END Wednesday 2025-01-15

The following Lemma is at the heart of the proof of Theorem 2.17.

Lemma 2.20. *With the hypotheses of Theorem 2.17, if $x \in \mathcal{X} \setminus \mathcal{M}$, then the conclusion of Theorem 2.17 holds with the subspace $\mathcal{M} + \mathbb{R}x$ in place of $\mathcal{X}$.*

Proof. For any $m_1, m_2 \in \mathcal{M}$, by hypothesis,

$$L(m_1) + L(m_2) = L(m_1 + m_2) \le p(m_1 + m_2) \le p(m_1 - x) + p(m_2 + x).$$

[2]There is also the Hahn-Banach Separation Theorem. Both theorems are often simple called *the* (sic) Hahn-Banach Theorem.

Rearranging gives, for $m_1, m_2 \in \mathcal{M}$,

$$L(m_1) - p(m_1 - x) \le p(m_2 + x) - L(m_2)$$

and thus

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

Now choose any real number λ satisfying

$$\sup_{m\in\mathcal{M}} \{L(m) - p(m-x)\} \le \lambda \le \inf_{m\in\mathcal{M}} \{p(m+x) - L(m)\}.$$

In particular, for $m \in \mathcal{M}$,

$$\begin{aligned} L(m) - \lambda &\le p(m-x) \\ L(m) + \lambda &\le p(m+x). \end{aligned} \tag{8}$$

Let $\mathcal{N} = \mathcal{M} + \mathbb{R}x$ and define $L' : \mathcal{N} \to \mathbb{R}$ by $L'(m + tx) = L(m) + t\lambda$ for $m \in \mathcal{M}$ and $t \in \mathbb{R}$. Thus L' is linear and agrees with L on $\mathcal{M}$ by definition. Moreover, by construction and equation 8,

$$\begin{aligned} L(m-x) &\le p(m-x) \\ L(m+x) &\le p(m+x). \end{aligned} \tag{9}$$

We now check that $L'(y) \le p(y)$ for all $y \in \mathcal{M} + \mathbb{R}x$. Accordingly, suppose $y \in \mathcal{N}$ so that there exists $m \in \mathcal{M}$ and $t \in \mathbb{R}$ such that $y = m + tx$. If $t = 0$ there is nothing to prove. If $t > 0$, then, in view of the second inequality of equation (9),

$$L'(y) = L'(m+tx) = t\left(L(\frac{m}{t}) + \lambda\right) \le t\,p(\frac{m}{t} + x) = p(m+tx) = p(y)$$

and a similar estimate, using the first inequality of equation (9), shows that

$$L'(m+tx) \le p(m+tx)$$

for $t < 0$. We have thus successfully extended L to a linear map $L' : \mathcal{N} \to \mathbb{R}$ satisfying $L'(n) \le p(n)$ for all $n \in \mathbb{N}$ and the proof is complete. $\square$

We make one further observation before turning to the proof of the Hahn-Banach Theorem. If T is a totally ordered set and $(\mathcal{N}_\alpha)_{\alpha\in T}$ are subspaces of a vector space $\mathcal{X}$ that are nested increasing in the sense that $\mathcal{N}_\alpha \subseteq \mathcal{N}_\beta$ for $\alpha \preceq \beta$, then $\mathcal{N} = \cup_{\alpha\in T}\mathcal{N}_\alpha$ is again a subspace of $\mathcal{X}$. By contrast, if $\mathcal{X}$ is a normed vector space and $\mathcal{N}_\alpha$ are (closed) subspaces of $\mathcal{X}$, then $\mathcal{N}$ will not necessarily be a (closed) subspace of $\mathcal{X}$.

Proof of Theorem 2.17. Let $\mathcal{L}$ denote the set of pairs $(L', \mathcal{N})$ where $\mathcal{N}$ is a subspace of $\mathcal{X}$ containing $\mathcal{M}$, and L' is an extension of L to $\mathcal{N}$ obeying $L'(y) \le p(y)$ on $\mathcal{N}$. Declare $(L'_1, \mathcal{N}_1) \preceq (L'_2, \mathcal{N}_2)$ if $\mathcal{N}_1 \subseteq \mathcal{N}_2$ and $L'_2|_{\mathcal{N}_1} = L'_1$. This relation $\preceq$ is a partial order on $\mathcal{L}$; that is $(\mathcal{L}, \preceq)$ is a partially ordered set. Further, Lemma 2.20 says if $(L', \mathcal{N})$ is maximal element, then $\mathcal{N} = \mathcal{X}$.

An exercise shows, given any increasing chain $(L'_\alpha, \mathcal{N}_\alpha)$ in $\mathcal{L}$ has as an upper bound $(L', \mathcal{N})$ in $\mathcal{L}$, where $\mathcal{N} := \bigcup_\alpha \mathcal{N}_\alpha$ and $L'(n_\alpha) := L'_\alpha(n_\alpha)$ for $n_\alpha \in \mathcal{N}_\alpha$. By Zorn's Lemma the collection $\mathcal{L}$ has a maximal element $(L', \mathcal{N})$ with respect to the order $\preceq$ and the proof is complete. □

The use of Zorn's Lemma in the proof of Theorem 2.17 is a typical - one knows how to carry out a construction one step a time, but there is no clear way to do it all at once. As an exercise, use Zorn's Lemma to prove that if V Is a vectors space and $S \subseteq V$ is a linearly independent set, then there is a basis B for V such that $B \supseteq S$.

In the special case that p is a seminorm, since $L(-x) = -L(x)$ and $p(-x) = p(x)$ the inequality $L \leq p$ is equivalent to $|L| \leq p$.

Corollary 2.21. *Suppose $\mathcal{X}$ is a real normed vector space, $\mathcal{M}$ is a subspace of $\mathcal{X}$, and L is a bounded linear functional on $\mathcal{M}$. If $C \geq 0$ and $|L(x)| \leq C\|x\|$ for all $x \in \mathcal{M}$, then there exists a bounded linear functional L' on $\mathcal{X}$ extending L such that $\|L'\| \leq C$.*

Proof. Apply the Hahn-Banach theorem with the Minkowski functional $p(x) = C\|x\|$. □

Before obtaining further corollaries, we extend Thoerem 2.17 to complex normed spaces. First, if $\mathcal{X}$ is a vector space over $\mathbb{C}$, then trivially it is also a vector space over $\mathbb{R}$, and there is a simple relationship between the $\mathbb{R}$- and $\mathbb{C}$-linear functionals.

Lemma 2.22. *Let $\mathcal{X}$ be a vector space over $\mathbb{C}$.*

(a) If $L : \mathcal{X} \to \mathbb{C}$ is a $\mathbb{C}$-linear functional, then $u(x) = \operatorname{real} L(x)$ defines an $\mathbb{R}$-linear functional on $\mathcal{X}$ and $L(x) = u(x) - iu(ix)$.
(b) Conversely, if $u : \mathcal{X} \to \mathbb{R}$ is $\mathbb{R}$-linear then $L(x) := u(x) - iu(ix)$ is $\mathbb{C}$-linear.
(c) If $L : \mathcal{X} \to \mathbb{C}$ is a $\mathbb{C}$-linear functional, $p : \mathcal{X} \to \mathbb{R}$ is a seminorm, and $u = \operatorname{real} L$,

then $|u(x)| \leq p(x)$ for all $x \in \mathcal{X}$ if and only if $|L(x)| \leq p(x)$ for all $x \in \mathcal{X}$.

Proof. Problem 2.5.

To prove the last statement, it is immediate that $|u(x)| \leq |L(x)|$ for all $x \in \mathcal{X}$.

Conversely, given x there is a unimodular α such that $\alpha L(x) = |L(x)|$. Hence,

$$|L(x)| = L(\alpha x) = |u(\alpha x)| \leq p(\alpha x) = |\alpha|\, p(x) = p(x).$$
□

Remark 2.23. Note that in passing from the real to the complex case, we must give up the generality of a Minkowski functional and instead content ourselves with the seminorm p.

END Friday 2025-01-17

Theorem 2.24 (The Hahn-Banach Theorem, complex version)**.** *Let $\mathcal{X}$ denote a complex vector space, p a seminorm on $\mathcal{X}$, and $\mathcal{M}$ a subspace of $\mathcal{X}$. If $L : \mathcal{M} \to \mathbb{C}$ is a $\mathbb{C}$-linear functional satisfying $|L(x)| \leq p(x)$ for all $x \in \mathcal{M}$, then there exists a $\mathbb{C}$-linear functional $L' : \mathcal{X} \to \mathbb{C}$ such that*

(i) $L'|_{\mathcal{M}} = L$ *and*
(ii) $|L'(x)| \le p(x)$ *for all* $x \in \mathcal{X}$.

Sketch of proof. Using the Lemma 2.22 and its notation: The proof consists of applying the real Hahn-Banach theorem (Corollary 2.21) to the $\mathbb{R}$-linear functional $u = \mathrm{Re}L$ to obtain a real linear functional $u' : \mathcal{X} \to \mathbb{R}$ extending u and satisfying $u'(x) \le p(x)$ for all $x \in \mathcal{X}$. The resulting complex functional L' associated to u' is then a desired extension of L. The details are left as an exercise. □

The following corollaries are quite important, and when the Hahn-Banach theorem is applied it is usually in one of the following forms:

Corollary 2.25. *Let* $\mathcal{X}$ *be a normed vector space over* $\mathbb{F}$ *(either* $\mathbb{R}$ *or* $\mathbb{C}$*).*

(i) *If* $\mathcal{M} \subseteq \mathcal{X}$ *is a subspace and* $L : \mathcal{M} \to \mathbb{F}$ *is a bounded linear functional, then there exists a bounded linear functional* $L' : \mathcal{X} \to \mathbb{F}$ *such that* $L'|_{\mathcal{M}} = L$ *and* $\|L'\| = \|L\|$.
(ii) (Linear functionals detect norms) *If* $x \in \mathcal{X}$ *is nonzero, there exists* $L \in \mathcal{X}^*$ *with* $\|L\| = 1$ *such that* $L(x) = \|x\|$.
(iii) (Linear functionals separate points) *If* $x \ne y$ *in* $\mathcal{X}$*, there exists* $L \in \mathcal{X}^*$ *such that* $L(x) \ne L(y)$.
(iv) (Linear functionals detect distance to subspaces) *If* $\mathcal{M} \subseteq \mathcal{X}$ *is a closed subspace and* $x \in \mathcal{X} \setminus \mathcal{M}$*, there exists* $L \in \mathcal{X}^*$ *such that*
  (a) $L|_{\mathcal{M}} = 0$;
  (b) $\|L\| = 1$; *and*
  (c) $L(x) = \mathit{dist}(x, \mathcal{M}) = \inf_{y \in \mathcal{M}} \|x - y\| > 0$.
(v) *if* $\mathcal{L}$ *is a linear submanifold of* $\mathcal{X}$ *and* $x \in \mathcal{X}$*, then* $x \in \overline{\mathcal{L}}$ *if and only if* $\lambda(x) = 0$ *for every* $\lambda \in \mathcal{X}^*$ *for which* $\mathcal{L} \subseteq \ker \lambda$.

Proof. To prove item (i) consider the (semi)norm $p(x) = \|L\|\,\|x\|$. By construction, $|L(x)| \le p(x)$ for $x \in \mathcal{M}$. Hence, there is a linear functional L' on $\mathcal{X}$ such that $L'|_{\mathcal{M}} = L$ and $|L'(x)| \le p(x)$ for all $x \in \mathcal{X}$. In particular, $\|L'\| \le \|L\|$. On the other hand, $\|L'\| \ge \|L\|$ since L' agrees with L on $\mathcal{M}$.

For item (ii), let $\mathcal{M}$ be the one-dimensional subspace of $\mathcal{X}$ spanned by x. Define a functional $L : \mathcal{M} \to \mathbb{F}$ by $L(t\frac{x}{\|x\|}) = t$. In particular, $|L(y)| = \|y\|$ for $y \in \mathcal{M}$ and thus $\|L\| = 1$. By (i), the functional L extends to a functional (still denoted L) on $\mathcal{X}$ such that $\|L\| = 1$.

An application of item (ii) to the vector $x - y$ proves item (iii).

To prove item (iv), let $\delta = \mathrm{dist}(x, \mathcal{M})$. Since $\mathcal{M}$ is closed, $\delta > 0$. Define a functional $L : \mathcal{M} + \mathbb{F}x \to \mathbb{F}$ by $L(y + tx) = t\delta$ for $y \in \mathcal{M}$ and $t \in \mathbb{F}$. Since for $t \ne 0$ and $y \in \mathcal{M}$,

$$\|y + tx\| = |t|\|t^{-1}y + x\| \ge |t|\delta = |L(y + tx)|,$$

by Hahn-Banach we can extend L to a functional $L \in \mathcal{X}^*$ with $\|L\| \le 1$.

Let $\mathcal{M} = \overline{\mathcal{L}}$. Thus $\mathcal{M}$ is a closed subspace of $\mathcal{X}$. If $\lambda \in \mathcal{X}^*$ and $\mathcal{L} \subseteq \ker\lambda$, then, by continuity, $\mathcal{M} \subseteq \ker\lambda$ proving one direction of item (v). For the remaining direction, apply item (iv) to $x \in \mathcal{X}\setminus\mathcal{M}$ to obtain a $\lambda \in \mathcal{X}^*$ such that $\mathcal{M} \subseteq \ker\lambda$, but $\lambda(x) \neq 0$. □

2.4. **The bidual and reflexive spaces.** Note that since $\mathcal{X}^*$ is a normed space, we can form its dual, denoted $\mathcal{X}^{**}$, and called the *bidual* or *double dual* of $\mathcal{X}$. There is a canonical relationship between $\mathcal{X}$ and $\mathcal{X}^{**}$. Each fixed $x \in \mathcal{X}$ gives rise to a linear functional $\hat{x} : \mathcal{X}^* \to \mathbb{F}$ via evaluation,

$$\hat{x}(L) := L(x).$$

Since $|\hat{x}(L)| = |L(x)| \leq \|L\|\,\|x\|$, the linear functional $\hat{x}$ is in $\mathcal{X}^{**}$ and $\|\hat{x}\| \leq \|x\|$.

Corollary 2.26. *(Embedding in the bidual) The map $x \to \hat{x}$ is an isometric linear map from $\mathcal{X}$ into $\mathcal{X}^{**}$.*

Proof. First, from the definition we see that

$$|\hat{x}(L)| = |L(x)| \leq \|L\|\|x\|$$

so $\hat{x} \in \mathcal{X}^{**}$ and $\|\hat{x}\| \leq \|x\|$. It is straightforward to check (recalling that the L's are linear) that the map $x \to \hat{x}$ is linear. Finally, to show that $\|\hat{x}\| = \|x\|$, fix a nonzero $x \in \mathcal{X}$. From Corollary 2.25(i) there exists $L \in \mathcal{X}^*$ with $\|L\| = 1$ and $L(x) = \|x\|$. But then for this x and L, we have $|\hat{x}(L)| = |L(x)| = \|x\|$ so $\|\hat{x}\| \geq \|x\|$, and the proof is complete. □

Definition 2.27. A Banach space $\mathcal{X}$ is called *reflexive* if the map $\hat{} : \mathcal{X} \to \mathcal{X}^{**}$ is surjective.

In other words, $\mathcal{X}$ is reflexive if the map $\hat{}$ is an (isometric) isomorphism of $\mathcal{X}$ with $\mathcal{X}^{**}$. For example, every finite dimensional Banach space is reflexive (Problem 2.6). Reflexive spaces often have nice properties. For instance, the distance from a point to a (closed) subspace is attained.

Needless to say, the proof of the Hahn-Banach theorem is thoroughly non-constructive, and in general it is an important (and often difficult) problem, given a normed space $\mathcal{X}$, to find some concrete description of the dual space $\mathcal{X}^*$. Usually doing so means finding a Banach space $\mathcal{Y}$ and a bounded (or, better, isometric) isomorphism $T : \mathcal{Y} \to \mathcal{X}^*$.

Example 2.28. Recall $\ell^1 = c_0^*$ isometrically and $\ell^\infty = (\ell^1)^*$ isometrically by Propositions 2.5 and 2.7 Moreover it is straightfoward to show that under the identification of Corollary 2.26, the canonical map $c_0 \to c_0^{**}$ corresponds to the natural inclusion of c_0 into ℓ^∞. Since c_0 is separable, but ℓ^∞ is not, c_0 is not reflexive.

Example 2.29 (Banach Limits)**.** The set

$$c = \left\{ f : \mathbb{N} \to \mathbb{F} \,\middle|\, \lim_{n\to\infty} f(n) \text{ exists} \right\},$$

is a subspace of ℓ^∞. The function $L : c \to \mathbb{F}$ defined by

$$L(f) = \lim_n f(n)$$

is a linear functional and it satisfies $|L(f)| \le \|f\|_\infty$. Hence L is continuous and $\|L\| \le 1$. On the other hand, letting $o : \mathbb{N} \to \mathbb{F}$ denote the function that is constantly equal to 1, we see $1 = |L(0)| = \|o\|$. Hence $\|L\| = 1$. Thus, by the Hahn-Banach Extension Theorem, L extends to a bounded linear functional on all of ℓ^∞ of norm 1. Any such extension of L is a *Banach limit*.

By example 2.6, elements $\lambda \in (\ell^1)^*$ are precisely of the form L_f for some $f \in \ell^\infty$, where $L_f(g) = \sum f(n)g(n)$ for $g \in \ell^1$. Thus $g \in \ell^1$ thus determines an element $\widehat{g}$ in $(\ell^1)^{**}$ by

$$\widehat{g}(f) = L_f(g).$$

Let $o_n : \mathbb{N} \to \mathbb{F}$ denote the function $o_n(j) = 0$ for $j \le n$ and $o_n(j) = 1$ for $j > n$ and observe,

$$\widehat{g}(o_n) = \sum_{j=n+1}^{\infty} g(n)$$

and $L(o_n) = 1$ for all n (where L is a Banach limit). It follows that $L \neq \widehat{g}$ and therefore the natural embedding of ℓ^1 into $(\ell^1)^{**} = (\ell^\infty)^*$ is not onto.

In fact more is true. Namely, there is no isometric isomorphism between ℓ^1 and $(\ell^\infty)^*$. As an outline of a proof, show, if $\mathcal{X}$ is a normed vector space and $\mathcal{X}^*$ is separable, then $\mathcal{X}$ is also separable. This fact, applied to ℓ^∞, shows $(\ell^\infty)^*$ is not separable. Since ℓ^1 is separable, the result follows. □

END Monday 2025-01-24

Remark 2.30. After we have studied the L^p and ℓ^p spaces in more detail, we will see that L^p is reflexive for $1 < p < \infty$.

We note in passing that if $\mathcal{X}$ is reflexive, then its dual $\mathcal{X}^*$ has a *unique predual*: that is, if $\mathcal{Y}$ is another Banach space and $\mathcal{Y}^*$ is isometrically isomorphic to $\mathcal{X}^*$, then in fact $\mathcal{Y}$ is isometrically isomorphic to $\mathcal{X}$. However this conclusion can fail when $\mathcal{X}$ is not reflexive; for example it turns out that ℓ^1 does not have a unique predual. See Problems 2.10 and 2.15. □

The embedding into the bidual has many applications; one of the most basic is the following.

Proposition 2.31 (Completion of normed spaces)**.** *If $\mathcal{X}$ is a normed vector space, then there is a Banach space $\overline{\mathcal{X}}$ and an isometric linear map $\iota : \mathcal{X} \to \overline{\mathcal{X}}$ such that the image $\iota(\mathcal{X})$ is dense in $\overline{\mathcal{X}}$.*

Proof. Embed $\mathcal{X}$ into $\mathcal{X}^{**}$ via the map $x \to \hat{x}$ and let $\overline{\mathcal{X}}$ be the closure of the image of $\mathcal{X}$ in $\mathcal{X}^{**}$. Since $\overline{\mathcal{X}}$ is a closed subspace of a complete space, it is complete. □

The space $\overline{\mathcal{X}}$ is called the *completion* of $\mathcal{X}$. It is unique in the sense that if $\mathcal{Y}$ is another Banach space and $j : \mathcal{X} \to \mathcal{Y}$ embeds $\mathcal{X}$ isometrically as a dense subspace of $\mathcal{Y}$, then $\mathcal{Y}$ is isometrically isomorphic to $\overline{\mathcal{X}}$. The proof of this fact is left as an exercise.

2.5. **Dual spaces and adjoint operators.** [Optional] Let $\mathcal{X}, \mathcal{Y}$ be normed spaces with duals $\mathcal{X}^*, \mathcal{Y}^*$. If $T : \mathcal{X} \to \mathcal{Y}$ is a linear transformation and $f : \mathcal{Y} \to \mathbb{F}$ is a linear functional, then $T^*f : \mathcal{X} \to \mathbb{F}$ defined by

$$(T^*f)(x) = f(Tx) \tag{10}$$

is a linear functional on $\mathcal{X}$. If T and f are both continuous (that is, bounded) then the composition T^*f is bounded, and more is true:

Theorem 2.32. *Let $T : \mathcal{X} \to \mathcal{Y}$ be a bounded linear transformation. The function $T^* : \mathcal{Y}^* \to \mathcal{X}^*$ defined, for $f \in \mathcal{Y}^*$, . Then:*

(i) For $f \in \mathcal{Y}^*$ the function T^*f defined by the formula (10) is in $\mathcal{X}^*$.
(ii) The mapping $T^* : \mathcal{Y}^* \to \mathcal{X}^*$ is a bounded linear map with $\|T^*\| = \|T\|$.

Proof. Since T is assumed bounded, for a fixed $f \in \mathcal{Y}^*$ and all $x \in \mathcal{X}$

$$|T^*f(x)| = |f(Tx)| \leq \|f\|\|Tx\| \leq \|f\|\|T\|\|x\|.$$

It follows that T^*f is bounded on $\mathcal{X}$ (thus, belongs to $\mathcal{X}^*$) and

$$\|T^*f\| \leq \|f\|\|T\|. \tag{11}$$

Thus T^* maps $\mathcal{Y}^*$ into $\mathcal{X}^*$ and it is straightforward to verify that T^* is linear. Moreover, the inequality of equation (11) also shows that T^* is bounded and $\|T^*\| \leq \|T\|$.

It remains to show $\|T^*\| \geq \|T\|$. Toward this end, let $0 < \epsilon < 1$ be given and choose $x \in \mathcal{X}$ with $\|x\| = 1$ and $\|Tx\| > (1 - \epsilon)\|T\|$. Now consider Tx. By the Hahn-Banach Theorem (Corollary 2.25(i)), there exists $f \in \mathcal{Y}^*$ such that $\|f\| = 1$ and $f(Tx) = \|Tx\|$. For this f,

$$\|T^*\| \geq \|T^*f\| \geq |T^*f(x)| = |f(Tx)| = \|Tx\| > (1 - \epsilon)\|T\|.$$

Hence, $\|T^*\| \geq (1 - \epsilon)\|T\|$. Since ϵ was arbitrary, $\|T^*\| \geq \|T\|$. □

END Monday 2025-01-27 also covered Proposition 2.12

2.6. **Duality for Sub and Quotient Spaces.** [Optional. Not covered Spring 2025] The Hahn-Banach Theorem allows for the identification of the duals of subspaces and quotients of Banach spaces. Informally, the dual of a subspace is a quotient and the dual of a quotient is a subspace. The precise results are stated below for complex scalars, but they hold also for real scalars.

Given a (closed) subspace $\mathcal{M}$ of the Banach space $\mathcal{X}$, let π denote the map from $\mathcal{X}$ to the quotient $\mathcal{X}/\mathcal{M}$. Recall (see Problem 1.22), the quotient is a Banach space with the norm,

$$\|z\| = \inf\{\|y\| : \pi(y) = z\}.$$

In particular, if $x \in \mathcal{X}$, then

$$\|\pi(x)\| = \inf\{\|x - m\| : m \in \mathcal{M}\}.$$

It is evident from the construction that π is continuous and $\|\pi\| \le 1$. Further, by Problem 1.20 (or see Proposition 2.33 below) if $\mathcal{M}$ is a proper (closed) subspace, then $\|\pi\| = 1$. In particular, $\pi^* : (\mathcal{X}/\mathcal{M})^* \to \mathcal{X}^*$ (defined by $\pi^*\lambda = \lambda \circ \pi$) is also continuous. Moreover, if $x \in \mathcal{M}$, then

$$\pi^*\lambda(x) = \lambda(\pi(x)) = 0.$$

Let

$$\mathcal{M}^\perp = \{f \in \mathcal{X}^* : f(x) = 0 \text{ for all } x \in \mathcal{M}\}.$$

($\mathcal{M}^\perp$ is called the *annihilator* of $\mathcal{M}$ in $\mathcal{X}^*$.) Recall, given $x \in \mathcal{X}$, the element $\hat{x} \in \mathcal{X}^{**}$ is defined by $\hat{x}(\tau) = \tau(x)$, for $\tau \in \mathcal{X}^*$. In particular,

$$\mathcal{M}^\perp = \cap_{x\in\mathcal{M}} \ker(\hat{x})$$

and thus $\mathcal{M}^\perp$ is a closed subspace of $\mathcal{X}^*$. Further, if $\lambda \in (\mathcal{X}/\mathcal{M})^*$, then $\pi^*\lambda \in \mathcal{M}^\perp$.

Proposition 2.33 (The dual of a quotient)**.** *The mapping $\psi : (\mathcal{X}/\mathcal{M})^* \to \mathcal{M}^\perp$ defined by*

$$\psi(\lambda) = \pi^*\lambda$$

is an isometric isomorphism; i.e., the mapping $\pi^* : (\mathcal{X}/\mathcal{M})^* \to \mathcal{X}^*$ is an isometric isomorphism onto $\mathcal{M}^\perp$.

Informally, the proposition is expressed as $(\mathcal{X}/\mathcal{M})^* = \mathcal{M}^\perp$.

Proof. The linearity of ψ follows from Theorem 2.32 as does $\|\psi\| = \|\pi\| \le 1$. To prove that ψ is isometric, let $\lambda \in (\mathcal{X}/\mathcal{M})^*$ be given. Automatically, $\|\psi(\lambda)\| \le \|\lambda\|$. To prove the reverse inequality, fix $r > 1$. Let $q \in \mathcal{X}/\mathcal{M}$ with $\|q\| = 1$ be given. There exists an $x \in X$ such that $\|x\| < r$ and $\pi(x) = q$. Hence,

$$|\lambda(q)| = |\lambda(\pi(x))\| = \|\psi(\lambda)(x)\| \le \|\psi(\lambda)\|\,\|x\| < r\|\psi(\lambda)\|.$$

Taking the supremum over such q shows $\|\lambda\| \le r\|\psi(\lambda)\|$. Finally, since $1 < r$ is arbitrary, $\|\lambda\| \le \|\psi(\lambda)\|$.

To prove that ψ is onto, and complete the proof, let $\tau \in \mathcal{M}^\perp$ be given. Fix $q \in \mathcal{X}/\mathcal{M}$. If $x, y \in \mathcal{X}$ and $\pi(x) = q = \pi(y)$, then $\tau(x) = \tau(y)$. Hence, the mapping $\lambda : \mathcal{X}/\mathcal{M} \to \mathbb{C}$ defined by $\lambda(q) = \tau(x)$ is well defined. That λ is linear is left as an exercise. To see that λ is continuous, observe that

$$|\lambda(q)| = |\tau(x)| \le \|\tau\|\,\|x\|,$$

for each $x \in \mathcal{X}$ such that $\pi(x) = q$. Taking the infimum over such x gives shows

$$|\lambda(q)| \leq \|\tau\| \, \|q\|.$$

Finally, by construction $\psi(\lambda) = \tau$. ☐

Since $\mathcal{M}^\perp$ is closed in $\mathcal{X}^*$, the quotient space $\mathcal{X}^*/\mathcal{M}^\perp$ is a Banach space. Let $\rho : \mathcal{X}^* \to \mathcal{X}^*/\mathcal{M}^\perp$ denote the quotient mapping. Suppose $\lambda \in \mathcal{M}^*$. By Corollary 2.25, there is an $f \in \mathcal{X}^*$ such that $f|_{\mathcal{M}} = \lambda$; that is f is a bounded extension of λ (and indeed f can be chosen such that $\|f\| = \|\lambda\|$). If f and g are two extensions of λ to bounded linear functionals on $\mathcal{X}^*$, then $f(x) - g(x) = 0$ for $x \in \mathcal{M}$. Hence $f - g \in \mathcal{M}^\perp$ or equivalently, $\rho(f) = \rho(g)$. Consequently, the mapping $\varphi : \mathcal{M}^* \to \mathcal{X}^*/\mathcal{M}^\perp$ defined by $\varphi(\lambda) = \rho(f)$ (where f is any bounded extension of λ to $\mathcal{X}$) is well defined. It is easily verified that φ is linear. Further, given $q \in \mathcal{X}^*/\mathcal{M}^\perp$, there is an $f \in \mathcal{X}^*$ such that $\rho(f) = q$. In particular, with $\lambda = f|_{\mathcal{M}}$ we have $\varphi(\lambda) = \rho(f)$. Therefore φ is onto.

Proposition 2.34 (The dual of a subspace)**.** *The mapping $\varphi : \mathcal{M}^* \to \mathcal{X}^*/\mathcal{M}^\perp$ is an isometric isomorphism.*

Proof. It remains to show that φ is an isometry, a fact that is an easy consequence of the Hahn-Banach Theorem. Fix $\lambda \in \mathcal{M}^*$ and let $q = \varphi(\lambda)$. If f is any bounded extension of λ to $\mathcal{X}^*$, then $\|f\| \geq \|\lambda\|$. Hence,

$$\begin{aligned}
\|\varphi(\lambda)\| &= \|q\| \\
&= \inf\{\|f\| : f \in \mathcal{X}^*, \quad \rho(f) = q\} \\
&= \inf\{\|f\| : f \in \mathcal{X}^*, \quad f|_{\mathcal{M}} = \lambda\} \\
&\geq \|\lambda\|.
\end{aligned}$$

On the other hand, by the Hahn-Banach Theorem there is a bounded extension g of λ with $\|g\| = \|\lambda\|$. Thus $\|\lambda\| \leq \|q\|$. ☐

A special case of the following useful fact was used in the proofs above. If $\mathcal{X}, \mathcal{Y}$ are vector spaces and $T : \mathcal{X} \to \mathcal{Y}$ is linear and $\mathcal{M}$ is a subspace of the kernel of T, then T induces a linear map $\widetilde{T} : \mathcal{X}/\mathcal{M} \to \mathcal{Y}$. A canonical choice is $\mathcal{M} = \ker(T)$ in which case $\widetilde{T}$ is one-one. If $\mathcal{X}$ is a Banach space, $\mathcal{Y}$ is a normed vector space and $\mathcal{M}$ is closed, then $\mathcal{X}/\mathcal{M}$ is a Banach space.

Lemma 2.35. *If $\mathcal{X}$ is a Banach space, $\mathcal{M}$ is a (closed) subspace of* $\ker(T)$*, $\mathcal{Y}$ is a normed vector space and $T : \mathcal{X} \to \mathcal{Y}$ is continuous, then the mapping $\widetilde{T}$ is bounded and $\|\widetilde{T}\| = \|T\|$.*

Proof. Let $\pi : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ denote the quotient map and observe that $\widetilde{T}\pi = T$. Since the quotient map π has norm 1 (see Problem 1.22), we see that $\|\widetilde{T}\| \leq \|T\|$. For the opposite inequality, let $0 < \epsilon < 1$ and choose $x \in \mathcal{X}$ such that $\|x\| = 1$ and $\|Tx\| > (1-\epsilon)\|T\|$. Then $\|\pi(x)\| \leq 1$ and

$$\|\tilde{T}\| \geq \|\tilde{T}\pi(x)\| = \|Tx\| > (1-\epsilon)\|T\|.$$

Letting ϵ go to zero finishes the proof. □

2.7. **Hahn-Banach separation theorems.** [Optional. Not covered Spring 2025]

Besides the extension theorem and its corollaries, the other important applications of the Hahn-Banach theorem consist of various separation theorems. We begin with a few definitions.

Definition 2.36. *Let $\mathcal{X}$ be a vector space. A* hyperplane *in $\mathcal{X}$ is a subspace $\mathcal{M}$ of codimension* 1*. An* affine hyperplane *is a set of the form $x + \mathcal{M} \subseteq X$, for some fixed $x \in \mathcal{X}$ and hyperplane $\mathcal{M}$.*

If $L : \mathcal{X} \to \mathbb{F}$ is a nonzero linear functional (bounded or not), the space $\mathcal{M} = \ker L$ is a hyperplane, and if we fix any scalar $t \in \mathbb{F}$ then the set $\{x \in \mathcal{X} : L(x) = t\}$ is an affine hyperplane. Conversely, any hyperplane is the kernel of a nonzero linear functional. (To see this, observe that if $\mathcal{M}$ is a hyperplane in $\mathcal{X}$, then, since it has codimension 1, for any fixed choice of a vector $y \in \mathcal{X} \setminus \mathcal{M}$ we can write every $x \in \mathcal{X}$ uniquely as $x = m + ty$ with $m \in \mathcal{M}$ and $t \in \mathbb{F}$. Then define $L(x) = t$.) Consequently, every affine hyperplane has the form $H = \{x \in \mathcal{X} : L(x) = t\}$ for some nonzero linear functional L and some scalar t.

Lemma 2.37. *If $\mathcal{M}$ is a hyperplane in a normed vector space $\mathcal{X}$, then $\mathcal{M}$ is either closed, or dense in $\mathcal{X}$.*

Proof. It is easy to check that the closure of subspace of $\mathcal{X}$ is again a subspace. It follows that $\overline{\mathcal{M}}$ is a subspace with $\mathcal{M} \subseteq \overline{\mathcal{M}} \subseteq \mathcal{X}$. Since $\mathcal{M}$ has codimension 1, we must have either $\overline{\mathcal{M}} = \mathcal{M}$ or $\overline{\mathcal{M}} = \mathcal{X}$. □

Proposition 2.38. *Let $\mathcal{X}$ be a normed vector space and $L : \mathcal{X} \to \mathbb{F}$ a linear functional. Then L is continuous (that is, bounded) if and only if $\ker L$ is closed. Consequenlty, L is continuous if and only if there exists a nonempty open set U such that $U \cap \ker L = \varnothing$.*

Proof. Trivially, if L is continuous then $\ker L$ is closed. Conversely, suppose $\mathcal{M} = \ker L$ is closed. We can then form the quotient space $\mathcal{X}/\mathcal{M}$, and since $\mathcal{M}$ is a hyperplane this space is one-dimensional. If we let π denote the quotient map and define $\tilde{L} : \mathcal{X}/\mathcal{M} \to \mathbb{F}$ by $\tilde{L}\pi = L$, then the linear functional $\tilde{L}$ is continuous (since its domain is finite-dimensional), and since the quotient map is also continuous, we conclude that L is continuous.

The second statement follows by combining the first statement with Lemma 2.37.

□

Recall that a set K in a real vector space $\mathcal{X}$ is called *convex* if for every $x, y \in K$ and every $0 \le t \le 1$, we have $tx + (1-t)y \in K$. Let $\mathcal{X}$ be a normed vector space over $\mathbb{R}$ and let $U \subseteq \mathcal{X}$ be a convex, open set containing 0. We define a function $p : \mathcal{X} \to [0, +\infty)$ by

$$p(x) = \inf\{r > 0 : \frac{1}{r}x \in U\}. \tag{12}$$

(To see that the definition makes sense, observe that since U is open and $0 \in U$, there exists $\delta > 0$ so that $x \in U$ whenever $\|x\| < \delta$. It follows that for every $x \in \mathcal{X}$, we have $\frac{1}{r}x \in U$ for all $r > \frac{\|x\|}{\delta}$; thus the set appearing in the definition is nonempty.) The functional p is sometimes called a *gauge* for the set U, it is important because of the following lemma.

Lemma 2.39. *Let $\mathcal{X}$ be a normed vector space over $\mathbb{R}$ and let $U \subseteq \mathcal{X}$ be a convex open set containing 0. Then the function p defined by (12) is a Minkowski functional, and $U = \{x \in \mathcal{X} : p(x) < 1\}$.*

Proof. If $r, s > 0$ then trivially $\frac{s}{r}x \in U$ if and only if $\frac{1}{r/s}x \in U$, and it follows that $p(sx) = sp(x)$ for all $s > 0$. Likewise it is immediate from the definition of p that $p(0) = 0$, so that $p(sx) = sp(x)$ for all $s \ge 0$. Next we show that $p(x) < 1$ if and only if $x \in U$: indeed, if x is in U then since U is open, there is a $\delta > 0$ such that $(1+\delta)x \in U$, thus $p(x) \le (1+\delta)^{-1} < 1$. On the other hand if $p(x) < 1$ then $\frac{1}{r}x \in U$ for some $0 < r < 1$, but then since U is convex and $0 \in U$, we can write $x = r(\frac{x}{r}) + (1-r) \cdot 0 \in U$.

Finally, let us show that $p(x+y) \le p(x) + p(y)$. Fix any $r, s > 0$ such that $\frac{x}{r}$ and $\frac{y}{s}$ belong to U. Since U is convex, the convex combination

$$\left(\frac{r}{r+s}\right)\frac{x}{r} + \left(\frac{s}{r+s}\right)\frac{y}{s} = \frac{x+y}{r+s}$$

belongs to U, so by what was just proved we have $p(\frac{x+y}{r+s}) < 1$. By homogeneity we conclude that $p(x+y) < r+s$, and finally by taking the infimum over r and s we get $p(x+y) \le p(x) + p(y)$.

□

Theorem 2.40 (Separation)**.** *Let $\mathcal{X}$ be a normed vector space over $\mathbb{R}$. If $U \subseteq \mathcal{X}$ is a nonempty, open, convex set, and $x \in \mathcal{X} \setminus U$, then there exists a bounded linear functional $L \in \mathcal{X}^*$ and a real number a such that $L(y) < a = L(x)$ for all $y \in U$.*

Proof. We first assume that $0 \in U$, the general case will follow by translation. Let $\mathcal{N}$ be the one-dimensional subspace $\mathbb{R}x$. Define L on $\mathcal{N}$ by putting $L(x) = 1$ and extending linearly. Let p be the gauge functional for U. Since $x \notin U$, we have $p(x) \ge 1$, so $1 = L(x) \le p(x)$. Since both L and p are positive homogeneous, we also have $L(tx) \le p(tx)$ for all $t \ge 0$. For $t < 0$, we have $L(tx) < 0 \le p(tx)$ (since $p \ge 0$ by definition). Thus, we have $L(y) \le p(y)$ for all y in the subspace $\mathcal{N}$. It follows from the Hahn-Banach theorem that there exists an extension of L to all of $\mathcal{X}$ (still denoted L)

such that $L(y) \leq p(y)$ for all $y \in \mathcal{X}$. It follows that $L(x) = 1$ and $L(y) \leq p(y) < 1$ for all $y \in U$. To see that this extension L is bounded, let $V = U - x$; then V is an open set in $\mathcal{X}$ and $L(y) < 0$ for all $y \in V$, that is, $V \cap \ker L = \varnothing$, so by the second part of Proposition 2.38 it follows that L is bounded.

Finally, in the case that U does not contain 0, we choose a point $x_0 \in U$ and apply the theorem to $U' := U - x_0$ and $x' = x - x_0$, to obtain a bounded functional such that $L(x) = 1 + L(x_0)$ and $L(y) < 1 + L(x_0)$ for all $y \in U$; the details are left to the reader. □

Lemma 2.41. i) *If $\mathcal{X}$ is a normed vector space over $\mathbb{R}$, $K \subseteq \mathcal{X}$ is a convex set, and x_0 is an interior point of K, then for every $x \in K$ and every $0 \leq t < 1$, the point $x_0 + t(x - x_0)$ is an interior point of K.*

ii) *If K is convex then int(K) is convex.*

iii) *If K is a closed, convex subset of $\mathcal{X}$ and K has nonempty interior, then K is equal to the closure of its interior.*

iv) *Let K be a closed, convex subset of $\mathcal{X}$ and suppose 0 is an interior point of K. If p is the gauge functional for the convex set $U =$ int(K), then $K = \{x \in \mathcal{X} | p(x) \leq 1\}$.*

Proof. For (i), by translation, there is no loss of generality in supposing that $x_0 = 0$. Fix $x \in K$, and fix δ such that $y \in K$ for all $\|y\| < \delta$. For $0 \leq t < 1$, put $\epsilon = (1-t)\delta$. If $\|z - tx\| < \epsilon$, then we can write $z = tx + y$ with $\|y\| < (1-t)\delta$. It follows that $\|(1-t)^{-1}y\| < \delta$, so $y' = (1-t)^{-1}y$ belongs to K. We have thus written $z = tx + (1-t)y'$ with $x, y' \in K$, so $z \in K$. That is, the open ball $B(tx, \epsilon)$ is contained in K.

(ii) follows immediately from (i).

For (iii), since $x = \lim(x_0 + t_n(x - x_0))$ for any sequence t_n increasing to 1, we see that every $x \in K$ is a limit of interior points.

For (iv), let $p(x) \leq 1$. If $p(x) < 1$ then $x \in U$ and thus $x \in K$. If $p(x) = 1$, then by the definiton of p we have $tx \in U$ for every $0 \leq t < 1$, so taking a sequence of scalars t_n increasing to 1, we get that x belongs to the closure of U so $x \in K$. Conversely, if $x \in K$, then by part (i) $tx \in U$ for every $0 \leq t < 1$, so $p(x) \leq 1$. □

Corollary 2.42 (Strict separation)**.** *Let $\mathcal{X}$ be a normed vector space over $\mathbb{R}$. If $K \subseteq \mathcal{X}$ is a closed, convex set with nonempty interior, and $x \in \mathcal{X} \backslash K$, then there exists a bounded linear functional $L \in \mathcal{X}^*$, and real numbers $a < b$ such that $L(y) \leq a < b = L(x)$ for all $y \in K$.*

Proof. Again we assume that 0 is an interior point of K, and leave the general case to the reader. Let $U = \text{int}(K)$ and let p be the gauge functional for U; by the lemma we have $K = \{x \in \mathcal{X} | p(x) \leq 1\}$. Thus, if $x \notin K$, then $p(x) > 1$. We may then choose a number $0 < t < 1$ so that $tx \notin K$. Applying the previous separation theorem to U and tx, we obtain a bounded linear functional L and a real number a such that $L(y) < a = L(tx)$

for all $y \in U$, we put $b := \frac{a}{t} = L(x)$. Since $L(y) < a$ on U, and L is continuous, and K is the closure of U (by item (iii) of the Lemma), we conclude that $L(y) \leq a$ for all $a \in K$, which completes the proof. □

2.8. **Problems.**

Problem 2.1. Prove, if $\mathcal{X}$ is any normed vector space, $\{x_1, \dots x_n\}$ is a linearly independent set in $\mathcal{X}$, and $\alpha_1, \dots \alpha_n$ are scalars, then there exists a bounded linear functional f on $\mathcal{X}$ such that $f(x_j) = \alpha_j$ for $j = 1, \dots n$. (Recall linear maps from a finite dimensional normed vector space to a normed vector space are bounded.)

Problem 2.2. Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{X} \to \mathcal{Y}$ a linear transformation. Prove T is bounded if and only if there exists a constant C such that for all $x \in \mathcal{X}$ and $f \in \mathcal{Y}^*$,

$$|f(Tx)| \leq C\|f\|\|x\|; \tag{13}$$

in which case $\|T\|$ is equal to the best possible C in (13).

Problem 2.3. Let $\mathcal{X}$ be a normed vector space. Show that if $\mathcal{M}$ is a closed subspace of $\mathcal{X}$ and $x \notin \mathcal{M}$, then $\mathcal{M} + \mathbb{F}x$ is closed. Use this result to give another proof that every finite-dimensional subspace of $\mathcal{X}$ is closed.

Problem 2.4. Prove, if $\mathcal{M}$ is a *finite-dimensional* subspace of a Banach space $\mathcal{X}$, then there exists a closed subspace $\mathcal{N} \subseteq \mathcal{X}$ such that $\mathcal{M} \cap \mathcal{N} = \{0\}$ and $\mathcal{M} + \mathcal{N} = \mathcal{X}$. (In other words, every $x \in \mathcal{X}$ can be written uniquely as $x = y + z$ with $y \in \mathcal{M}$, $z \in \mathcal{N}$.) *Hint:* Choose a basis $x_1, \dots x_n$ for $\mathcal{M}$ and construct, using Problem 2.1 and the Hahn-Banach Theorem, bounded linear functionals $f_1, \dots f_n$ on $\mathcal{X}$ such that $f_i(x_j) = \delta_{ij}$. Now let $\mathcal{N} = \cap_{i=1}^n \ker f_i$. (Warning: this conclusion can fail badly if $\mathcal{M}$ is not assumed finite dimensional, even if $\mathcal{M}$ is still assumed closed. Perhaps the first known example is that c_0 is not *complemented* in ℓ^∞, though it is nontrivial to prove.)

Problem 2.5. Prove Proposition 2.22.

Problem 2.6. Prove every finite-dimensional Banach space is reflexive.

Problem 2.7. Let B denote the subset of ℓ^∞ consisting of sequences which take values in $\{-1, 1\}$. Show that any two (distinct) points of B are a distance 2 apart. Show, if C is a countable subset of ℓ^∞, then there exists a $b \in B$ such that $\|b - c\| \geq 1$ for all $c \in C$. Conclude ℓ^∞ is not separable. Prove there is no isometric isomorphism $\Lambda : c_0 \to \ell^\infty$.

Problem 2.8. [This problem belongs in the section with signed measures] Prove, if μ is a finite regular (signed) Borel measure on a compact Hausdorff space, then the linear function $L_\mu : C(X) \to \mathbb{R}$ defined by

$$L_\mu(f) = \int_X f \, d\mu$$

is bounded (continuous) and $\|L_\mu\| = \|\mu\| := |\mu|(X)$. (See the Riesz-Markov Theorem for positive linear functionals.)

Problem 2.9. Let $\mathcal{X}$ and $\mathcal{Y}$ be normed vector spaces and $T \in L(\mathcal{X}, \mathcal{Y})$.

a) Consider $T^{**} : \mathcal{X}^{**} \to \mathcal{Y}^{**}$. Identifying $\mathcal{X}, \mathcal{Y}$ with their images in $\mathcal{X}^{**}$ and $\mathcal{Y}^{**}$, show that $T^{**}|_{\mathcal{X}} = T$.
b) Prove T^* is injective if and only if the range of T is dense in $\mathcal{Y}$.
c) Prove that if the range of T^* is dense in $\mathcal{X}^*$, then T is injective; if $\mathcal{X}$ is reflexive then the converse is true.
d) Assuming now that $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, prove that $T : \mathcal{X} \to \mathcal{Y}$ is invertible if and only if T^* is invertible, in which case $(T^*)^{-1} = (T^{-1})^*$.

Problem 2.10. a) Prove that if $\mathcal{X}$ is reflexive, then $\mathcal{X}^*$ is reflexive. (Hint: let $\iota : \mathcal{X} \to \mathcal{X}^{**}$ be the canonical inclusion; by assumption ι is invertible. Compute $(\iota^{-1})^*$.)
b) Prove that if $\mathcal{X}$ is reflexive and $\mathcal{M} \subseteq \mathcal{X}$ is a closed subspace, then $\mathcal{M}$ is reflexive.
c) Prove that a Banach space $\mathcal{X}$ is reflexive if and only if $\mathcal{X}^*$ is reflexive.
d) Prove that if $\mathcal{X}$ is reflexive and $\mathcal{Y}$ is another Banach space with $\mathcal{Y}^*$ isometrically isomorphic to $\mathcal{X}^*$, then $\mathcal{Y}$ is isometrically isomorphic to $\mathcal{X}$. (This conclusion can fail if $\mathcal{X}$ is not reflexive; see Problem 2.15.)

Problem 2.11. Prove, if $\mathcal{X}$ is a Banach space and $\mathcal{X}^*$ is separable, then $\mathcal{X}$ is separable. [*Hint:* let $\{f_n\}$ be a countable dense subset of $\mathcal{X}^*$. For each n choose x_n such that $\|x_n\| = 1$ and $|f_n(x_n)| \geq \frac{1}{2}\|f_n\|$. Show that the set of $\mathbb{Q}$-linear combinations of $\{x_n\}$ is dense in $\mathcal{X}$.]

Problem 2.12. a) Prove there exists a bounded linear functional $L \in (\ell^\infty)^*$ with the following property: whenever $f \in \ell^\infty$ and $\lim_{n\to\infty} f(n)$ exists, then $L(f)$ is equal to this limit. (Hint: first show that the set of such f forms a subspace $\mathcal{M} \subseteq \ell^\infty$). Such an L is a *Banach limit*.
b) Show that such a functional L is not equal to L_g for any $g \in \ell^1$; thus the map $T : \ell^1 \to (\ell^\infty)^*$ given by $T(g) = L_g$ is not surjective.
c) Give another proof that T is not surjective, using Problem 2.11.

Problem 2.13. Let $\mathcal{X}$ be a normed space and let $K \subseteq \mathcal{X}$ be a convex set. (Recall, this means that whenever $x, y \in K$, then $\frac{1}{2}(x + y) \in K$; equivalently, $tx + (1 - t)y \in K$ for all $0 \leq t \leq 1$.) A point $x \in K$ is called an *extreme point* of K whenever $y, z \in K$, $0 < t < 1$, and $x = ty + (1-t)z$, then $y = z = x$. (That is, the only way to write x as a convex combination of elements of K is the trivial way.)

a) Let $\mathcal{X}$ be a normed space and let $B = ball(\mathcal{X})$ denote the (closed) unit ball of $\mathcal{X}$. Prove that $x \in B$ is *not* an extreme point of B if and only if there exists a nonzero $y \in B$ such that $\|x \pm y\| \leq 1$.

b) Prove that if $\mathcal{X}$ and $\mathcal{Y}$ are normed spaces, and $T : \mathcal{X} \to \mathcal{Y}$ is a surjective linear isometry, (so that $\mathcal{X}$ and $\mathcal{Y}$ are isometrically isomorphic) then T induces a bijection between the extreme points of $ball(\mathcal{X})$ and $ball(\mathcal{Y})$.

c) Let ℓ^p_n denote the (real) Banach space $\mathbb{R}^n$ equipped with the ℓ^p norm, $1 \le p \le \infty$. Prove that ℓ^1_2 and ℓ^∞_2 are isometrically isomorphic, but that there is *no* isometry between ℓ^1_3 and ℓ^∞_3.

Problem 2.14. a) Show that the extreme points of the unit ball of ℓ^1 are precisely the points of the form λe_n where $|\lambda| = 1$ and e_n is the sequence which is 1 in the n^{th} entry and 0 elsewhere. (See Problem 2.13).

b) Determine the extreme points of the unit ball of ℓ^∞.

c) Show that the unit ball of c_0 has *no* extreme points.

Problem 2.15. Let

$$c = \left\{ f : \mathbb{N} \to \mathbb{F} \,\middle|\, \lim_{n\to\infty} f(n) \text{ exists} \right\},$$

and equip c with the supremum norm $\|f\|_\infty := \sup |f(n)|$.

a) Show that $c^* \cong \ell^1$ isometrically.

b) Prove that c is boundedly isomorphic to c_0.

c) Prove that c is not *isometrically* isomorphic to c_0. (Hint: examine the extreme points of the unit balls of c and c_0; see Problems 2.13 and 2.14 .)

(This problem provides an example of Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ such that $\mathcal{X}$ and $\mathcal{Y}$ are not isometrically isomorphic, but $\mathcal{X}^*$ and $\mathcal{Y}^*$ are. So in general we cannot recover $\mathcal{X}$ (isometrically) from $\mathcal{X}^*$. In fact the situation is worse, ℓ^1 has isometric preduals which are not even boundedly isomorphic to c_0, but the construction is more involved and outside the scope of these notes.)

3. The Baire Category Theorem and applications

This section contains three important applications of the Baire category theorem in functional analysis. These are the Principle of Uniform boundedness (also known as the Banach-Steinhaus theorem), the Open Mapping Theorem, and the Closed Graph Theorem. (In learning these theorems, keep careful track of what completeness hypotheses are needed.)

3.1. **Baire's Theorem.** Recall, a set D in a metric space X is *dense* (in X) if $\overline{D} = X$. Lemma 3.1 below should be familiar. We will use the notation $B(x, r)$, for the open ball of radius $r > 0$ center to the point x in a metric space $X = (X, d)$

$$B(x, r) = \{y \in X : d(x, y) < r\};$$

and F° for the *interior* of a subset F of a metric space X.

Lemma 3.1. *Suppose X is a metric space.*

(a) For a subset $D \subseteq X$ of X, the following are equivlalent:

(i) D is dense in X;

(ii) D^c does not contain a nonempty open set $((D^c)^\circ = \varnothing$;

(iii) if $\varnothing \neq U$ is open, then $D \cap U \neq \varnothing$.

(b) If $U \subseteq X$ is open and $x \in U$, then there is an $r > 0$ such that $\overline{B(x,r)} \subseteq U$.

(c) A subset F of X is closed with empty interior if and only if F^c is open and dense.

Theorem 3.2 (The Baire Category Theorem)**.** *Suppose X is a complete metric space.*

(a) If $(U_n)_{n=1}^\infty$ is a sequence of open dense subsets of X, then $\cap_{n=1}^\infty U_n \neq \emptyset$.

(b) If $(F_n)_n$ is a sequence of closed sets with empty interior, then $\cup F_n \neq X$.

Remark 3.3. We will actually prove that $\cap U_n$ is dense in X. This conclusion is in fact equivalent to the conclusion that $\cap U_n \neq \varnothing$.

Theorem 3.2 is true if X is a locally compact Hausdorff space and there are connections between the Baire Category Theorem and the axiom of choice. □

The following lemma should be familiar from advanced calculus. It will be used in the proof of Theorem 3.2.

Lemma 3.4. *Let X be a complete metric space and suppose (C_n) is a sequence of subsets of X. If*

(i) each C_n is nonempty;
(ii) (C_n) is nested decreasing;
(iii) each C_n is closed; and
(iv) $(diam(C_n))$ converges to 0,

then there is an $x \in X$ such that

$$\{x\} = \cap C_n.$$

Moreover, if (x_n) is a sequence from X and $x_n \in C_n$ for each n, then (x_n) converges to some x.

Proof of Theorem 3.2. The two items are equivalent, but for our purposes it is enough to show item (a) implies item (b). To this end, suppose item (a) holds, (F_n) is a sequence of closed sets such that $F_n^\circ = \varnothing$ for all n and $X = \cup F_n$. Taking complements, $\varnothing = \cap F_n^c$. By Lemma 3.1, the sets F_n^c are open and dense in X by Lemma 3.1. Hence X is not complete and therefore item (a) implies item (b). Thus it suffices to prove item (a).

To prove item (a), let $(U_n)_{n=1}^\infty$ be a sequence of open dense sets in X and let $I = \cap U_n$. To prove I is dense, it suffices to show that I has nontrivial intersection with every nonempty open set W by Lemma 3.1. Fix such a W. Since, by Lemma 3.1, U_1 is dense, there is a point $x_1 \in W \cap U_1$. Since U_1 and W are open, there is a radius $0 < r_1 < 1$ such that the $\overline{B(x_1, r_1)}$ is contained in $W \cap U_1$ by Lemma 3.1. Similarly, since U_2 is dense and open there is a point $x_2 \in B(x_1, r_1) \cap U_2$ and a radius $0 < r_2 < \frac{1}{2}$ such that

$$\overline{B(x_2, r_2)} \subseteq B(x_1, r_1) \cap U_2 \subseteq W \cap U_1 \cap U_2.$$

Continuing inductively, since each U_n is dense and open there is a sequence of points $(x_n)_{n=1}^{\infty}$ and radii $0 < r_n < \frac{1}{n}$ such that

$$\overline{B(x_n, r_n)} \subseteq B(x_{n-1}, r_{n-1}) \cap U_n \subseteq W \cap (\cap_{j=1}^{n} U_n).$$

The sequence of sets $(\overline{B(x_n, r_n)})$ satisfies the hypothesis of Lemma 3.4 and X is complete. Hence there is an $x \in X$ such that

$$x \in \cap_n \overline{B(x_n, r_n)} \subseteq W \cap I. \qquad \square$$

3.2. **Category.** Baire's theorem is used as a kind of pigeonhole principle: the "thick" complete metric space X cannot be expressed as a countable union of "thin" closed sets without interior.

Definition 3.5. A subset E of a metric space X is *nowhere dense* (in X) if its closure has empty interior; that is $(\overline{E})^{\circ} = \varnothing$.

A set F in a metric space X is *first category* (or *meager*) if it can be expressed as a countable union of nowhere dense sets. In particular, a countable union of first category sets is first category.

A set G is *second category* if it is not first category.

Corollary 3.6 (The Baire Category Theorem restated)**.** *If X is a complete metric space, then X is not a countable union of nowhere dense sets; that is, X is of second category in itself.*

Proof. Suppose $X = \cup_{n=1}^{\infty} E_n$ where each E_n is nowhere dense. It follows that $X = \cup_{n=1}^{\infty} F_n$, where each $F_n = \overline{E_n}$ is closed and with empty interior. Hence, by Theorem 3.2 item (b), X is not complete. $\square$

Corollary 3.7. *An infinite dimensional Banach space can not have a countable basis. In particular, there is no norm on c_{00} that makes it a Banach space; ditto for the vector space of polynomials.*

Proof. The proof is left as an exercise. As a starting point, show, if $\mathcal{M}$ is finite dimensional subspace of an infinite dimensional Banach space $\mathcal{X}$, then $\mathcal{M}$ is nowhere dense in $\mathcal{X}$. $\square$

3.3. **The Principle of Uniform Boundedness.**

Theorem 3.8 (The Principle of Uniform Boundedness (PUB))**.** *Suppose $\mathcal{X}, \mathcal{Y}$ are normed spaces and $\{T_\alpha : \alpha \in A\} \subseteq B(\mathcal{X}, \mathcal{Y})$ is a collection of bounded linear transformations from $\mathcal{X}$ to $\mathcal{Y}$. Let B denote the set*

$$B := \{x \in X : \sup_\alpha \|T_\alpha x\| < \infty\}. \tag{14}$$

If B is of the second category (thus not a countable union of nowhere dense sets) in X, then

$$\sup_\alpha \|T_\alpha\| < \infty.$$

In particular, if $\mathcal{X}$ is complete and if the collection $\{T_\alpha : \alpha \in A\}$ is pointwise bounded, then it is uniformly bounded.

Proof. For notational convenience, set $M(x) = \sup_\alpha \|T_\alpha x\| < \infty$.

For each integer $n \geq 1$ consider the set

$$V_n := \{x \in \mathcal{X} : M(x) > n\}.$$

Since each T_α is continuous (bounded), the sets V_n are open. (Indeed, for each α the map $x \to \|T_\alpha x\|$ is continuous from $\mathcal{X}$ to $\mathbb{R}$, so if $\|T_\alpha x\| > n$ for some α then also $\|T_\alpha y\| > n$ for all y sufficiently close to x.) Let E_n denote the complement of V_n and observe that $B = \cup_{n=1}^\infty E_n$. Since B is assumed to be of the second category, there is an N such that $(\overline{E_N})^\circ$ is not empty. Since E_N is closed, it follows that E_N has nonempty interior; i.e., there is an $x_0 \in E_N$ and $r > 0$ so that $B(x_0, r) \subseteq E_N$. α and every $\|x\| < r$, expressing $x = (x - x_0) + x_0$ as the sum of two elements of $B(x_0, r)$ gives

$$\|T_\alpha x\| \leq \|T_\alpha(x - x_0)\| + \|T_\alpha x_0\| \leq N + N.$$

That is, if $\|x\| < r$, then $M(x) \leq 2N$. By rescaling we conclude that if $\|x\| < 1$, then $\|T_\alpha x\| \leq 2N/r$ for all α and thus $\sup_\alpha \|T_\alpha\| \leq 2N/r < \infty$. □

The following result is one of the many corollaries to the PUB.

Corollary 3.9. *Suppose $\mathcal{X}$ is a Banach space and $\mathcal{Y}$ is a normed vector space. If (T_n) is a sequence of bounded operators $T_n : \mathcal{X} \to \mathcal{Y}$ that converges pointwise to a (necessarily linear) map $T : \mathcal{X} \to \mathcal{Y}$, then T is bounded.*

Outline of proof. In a metric space, convergent sequences are bounded. Hence $(T_n x)_n$ is bounded in $\mathcal{Y}$ for each $x \in \mathcal{X}$. Thus the set $\mathcal{X} = \{x \in \mathcal{X} : \sup\{\|T_n x\| : n \in \mathbb{N}\} < \infty\}$ is of second category in $\mathcal{X}$. Thus $C = \sup\{\|T_n\| : n \in \mathbb{N}\} < \infty$. Thus the proof reduces to showing $\|Tx\| \leq C\|x\|$ for all $x \in \mathcal{X}$, a task that is left as an exercise for the gentle reader. □

3.4. **Open mapping.** Given a subset B of a vector space $\mathcal{X}$ and a scalar $s \in \mathbb{F}$, let $sB = \{sb : b \in B\}$. Similarly, for $x \in \mathcal{X}$, let $B - x = \{b - x : b \in B\}$. Let $\mathcal{X}, \mathcal{Y}$ be normed vector spaces and suppose $T : \mathcal{X} \to \mathcal{Y}$ is linear. If $B \subseteq \mathcal{X}$ and $s \in \mathbb{F}$ is nonzero, then $T(sB) = sT(B)$ and further, an easy argument shows $\overline{T(sB)} = s\,\overline{T(B)}$. It is also immediate that if B is open, then so is $B - x$.

Recall that if X, Y are topological spaces, a mapping $f : X \to Y$ is called *open* if $f(U)$ is open in Y whenever U is open in X. In particular, if f is a bijection, then f is open if and only if f^{-1} is continuous. In the case of normed linear spaces the condition that a linear map is open is refined by the Open Mapping Theorem, Theorem 3.11 below. But first a lemma.

Lemma 3.10 (Translation and Dilation lemma)**.** *Let $\mathcal{X}, \mathcal{Y}$ be normed vector spaces, let B denote the open unit ball of $\mathcal{X}$, and let $T : \mathcal{X} \to \mathcal{Y}$ be a linear map. The following are equivalent.*

(i) The map T is open;
(ii) $T(B)$ contains an open ball centered at 0;
(iii) there is an $s > 0$ such that $T(sB)$ contains an open ball centered at 0; and
(iv) $T(sB)$ contains an open ball centered at 0 for each $s > 0$.

In the proof of this lemma and that of Theorem 3.11 to follow, we use $B^{\mathcal{X}}(x, r)$ and $B^{\mathcal{Y}}(y, s)$ to denote the open balls centered to x and y with radii r and s in $\mathcal{X}$ and $\mathcal{Y}$ respectively when needed to avoid ambiguity.

Proof. This result is more or less immediate from the fact that, for fixed $z_0 \in \mathcal{X}$ and $r \in \mathbb{F}$, the translation map $z \to z + z_0$ and the dilation map $z \to rz$ are continuous in a normed vector space.

The implication item (i) implies item (ii) is immediate. The fact that $T(sB) = sT(B)$ for $s > 0$ readily shows items (ii), (iii) and (iv) are all equivalent.

To finish the proof it suffice to show item (iv) implies item (i). Accordingly, suppose item (iv) holds and let $U \subseteq X$ be a given open set. To prove that $T(U)$ is open, let $y \in T(U)$ be given. There is an $x \in U$ such that $T(x) = y$. There is an $s > 0$ such that the ball $B(x, s)$ lies in U; that is $B(x, s) \subseteq U$. The ball $B(0, s) = B(x, s) - x$ is an open ball centered to 0. By hypothesis there is an $r > 0$ such that $B^{\mathcal{Y}}(0, r) \subseteq T(B(0, s))$. By linearity of T,

$$\begin{aligned} B^{\mathcal{Y}}(y, r) =& B^{\mathcal{Y}}(0, r) + y \subseteq T(B(0, s)) + y \\ =& T(B(0, s)) + T(x) = T(B(0, s) + x) = T(B(x, s)) \subseteq T(U). \end{aligned}$$

Thus $T(U)$ is open and the proof is complete. $\square$

Theorem 3.11 (Open Mapping)**.** *Suppose that $\mathcal{X}$ is a Banach space, $\mathcal{Y}$ is a normed vector space and $T : \mathcal{X} \to \mathcal{Y}$ is bounded. If the range of T is of second category, then*

(i) $T(\mathcal{X}) = \mathcal{Y}$*;*
(ii) $\mathcal{Y}$ *is complete (so a Banach space); and*
(iii) T *is open.*

In particular, if $\mathcal{X}, \mathcal{Y}$ *are Banach spaces, and* $T : \mathcal{X} \to \mathcal{Y}$ *is bounded and onto, then* T *is an open map.*

Proof. Assuming T is open and letting B denote the open unit ball in $\mathcal{X}$, by Lemma 3.10, there is an $r > 0$ such that $B^{\mathcal{Y}}(0, r) \subseteq T(B)$. Hence,

$$\mathcal{Y} = \cup_{n=1}^{\infty} B^{\mathcal{Y}}(0, n\, r) \subseteq T(nB) \subseteq T(\mathcal{X}),$$

so that item (i) holds. (Here the superscript $\mathcal{Y}$ is used to emphasize this ball is in $\mathcal{Y}$.) That is, item (iii) implies item (i).

To prove item (iii), let $B(x, r)$ denote the open ball of radius r centered at x in $\mathcal{X}$. Trivially $\mathcal{X} = \bigcup_{n=1}^{\infty} B(0, n)$ and thus $\operatorname{range} T = T(\mathcal{X}) = \bigcup_{n=1}^{\infty} T(B(0, n))$. Since the range of T is assumed second category, there is an N such that $T(B(0, N))$ is second category and hence somewhere dense. In other words, $\overline{T(B(0, N))}$ has nonempty interior. By scaling (see Lemma (3.10)), $\overline{T(B(0, 1))}$ has nonempty interior. Hence, there exists $p \in \mathcal{Y}$ and $r > 0$ such that $\overline{T(B(0, 1))}$ contains the open ball $B^{\mathcal{Y}}(p, r)$. It follows that for all $\|y\| < r$,

$$y = (y + p) - p \in \overline{T(B(0, 1))} + \overline{T(B(0, 1))} \subseteq \overline{T(B(0, 2))},$$

where we have used $-T(B(0, 1)) = T(B(0, 1))$. In other words,

$$B^{\mathcal{Y}}(0, r) \subseteq \overline{T(B(0, 2))}.$$

By scaling, it follows that, for $n \in \mathbb{N}$,

$$B^{\mathcal{Y}}(0, \frac{r}{2^{n+1}}) \subseteq \overline{T(B(0, \frac{1}{2^n}))}.$$

END Friday 2025-01-31

We will use the hypothesis that $\mathcal{X}$ is complete to prove $B^{\mathcal{Y}}(0, \frac{r}{4}) \subseteq T(B(0, 1))$, which, by Lemma 3.10, implies T is open. Accordingly let y such that $\|y\| < \frac{r}{4}$ be given. Since y is in the closure of $T(B(0, \frac{1}{2}))$, there is a $y_1 \in T(B(0, \frac{1}{2}))$ such that $\|y - y_1\| < \frac{r}{8}$. Since $y - y_1 \in B^{\mathcal{Y}}(0, \frac{r}{8})$ it is is in the closure of $T(B(0, \frac{1}{4}))$. Thus there is a $y_2 \in T(B(0, \frac{1}{4}))$ such that $\|(y - y_1) - y_2\| < \frac{r}{16}$. Continuing in this fashion produces a sequence $(y_j)_{j=1}^{\infty}$ from $\mathcal{Y}$ such that,

(a) $\|y - \sum_{j=1}^{n} y_j\| \le \frac{r}{2^{n+2}}$; and
(b) $y_n \in T(B(0, \frac{1}{2^n}))$

for all n. It follows the sequence $(s_m = \sum_{j=1}^{m} y_j)_m$ converges to y. Further, for each j there is an $x_j \in B(0, \frac{1}{2^j})$ such that $y_j = Tx_j$. Thus, setting $t_m = \sum_{j=1}^{m} x_j$, we have

$Ts_m = t_m$ and

$$\sum_{j=1}^{\infty} \|x_j\| < \sum_{k=1}^{\infty} 2^{-k} = 1,$$

thus the sequence $(t_m)_m$ converges to some x with $\|x\| \leq \sum_{j=1}^{\infty} \|x_j\| < 1$, that is, $x \in B(0,1)$. It follows that $y = Tx$ by continuity of T. Consequently $y \in T(B(0,1))$ and the proof of item (iii) is complete.

To prove item (ii), let $\mathcal{M}$ denote the kernel of T and $\widetilde{T}$ the mapping $\widetilde{T} : \mathcal{X}/\mathcal{M} \to \mathcal{Y}$ determined by $\widetilde{T}\pi = T$; that is $\widetilde{T}(\pi(x)) = T(x)$ for $x \in \mathcal{X}$. By construction $\widetilde{T}$ is one-one and by Lemma 2.35, it is continuous. Further its range is the same as the range of T, namely $\mathcal{Y}$, and is thus second category. Hence, by what has already been proved, $\widetilde{T}$ is an open map. and consequently $\widetilde{T}^{-1}$ is continuous. Hence $\mathcal{X}/\mathcal{M}$ and $\mathcal{Y}$ are isomorphic (though of course not necessarily isometrically isomorphic) as normed vector spaces. Therefore, since $\mathcal{X}/\mathcal{M}$ is complete (Proposition 1.29), so is $\mathcal{Y}$. □

Note that the proof of item (ii) in the Open Mapping Theorem shows, in the case that in the case that T is one-one and its range is of second category, that T is onto and its inverse is continuous. In particular, if $T : \mathcal{X} \to \mathcal{Y}$ is a continuous bijection and $\mathcal{Y}$ is a Banach space (so the range of T is second category), then T^{-1} is continuous.

Corollary 3.12 (The Banach Isomorphism Theorem)**.** *If $\mathcal{X}, \mathcal{Y}$ are Banach spaces and $T : \mathcal{X} \to \mathcal{Y}$ is a bounded bijection, then T^{-1} is also bounded (hence, T is an isomorphism).*

Proof sketch. Note that when T is bijection, T is open if and only if T^{-1} is continuous. The result thus follows from the Open Mapping Theorem and Proposition 1.32. □

The following examples show that the hypothesis that $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, and not just normed vector spaces, is needed in Corollary 3.12.

Example 3.13. This example shows that the assumption that the range of T is second category in $\mathcal{Y}$ is necessary in Theorem 3.11.

Let $\mathcal{X}$ denote the Banach space ℓ^1 and let $\mathcal{Y} = \ell^1$ as a linear manifold in c_0 with the c_0 (sup) norm. So $\mathcal{Y}$ is a normed space, but not a Banach space. The identity map $\iota \mathcal{X} to \mathcal{Y}$ is a bijection. It is also continuous since the supremum norm of an element of ℓ^1 dominates its ℓ^1 norm. Let $G = B^{\mathcal{X}}(0,1) \subseteq \mathcal{X}$ denote the (open) unit ball in X. Thus G is open in $\mathcal{X}$. Given $r > 0$ choose $n \in \mathbb{N}$ such that $n > \frac{2}{r}$ and let $x = \frac{r}{2} \sum_{j=1}^{n} e_j$, where $e_j \in \ell^1$ is the function $e_j : \mathbb{N} \to \mathbb{F}$ defined by

$$e_j(m) = \begin{cases} 1 & \text{if } m = j \\ 0 & \text{if } m \neq j. \end{cases}$$

Observe that $\|x\|_\infty < r$, but $\|x\|_1 > 1$. Hence $x \in B^{\mathcal{Y}}(0,r)$, but $x \notin G$. Thus $B^{\mathcal{Y}}(0,r) \not\subseteq G$ for any choice of $r > 0$, which means 0 is not an interior point of G (in $\mathcal{Y}$). Hence G is not open in $\mathcal{Y}$.

The argument above of course shows that the range of ι is not second category in $\mathcal{Y}$. Here is simple direct proof of this fact. First note that the sequence $g_n = \frac{1}{n}\sum_{j=1}^n e_j$ converges to 0 in c_0 and $g_n \in \ell^1$ with $\|g_n\|_1 = 1$. For $N \in \mathbb{N}$, let $B_N = \{f \in \ell^1 : \|f\|_1 \leq N\} \subseteq \mathcal{Y}$. Verify that each B_N is closed in $\mathcal{Y}$. On the other hand, given N, the sequence $f_k = f + 3Ng_k$ converges to f in c_0 and so in $\mathcal{Y}$, but $\|f + 3Ng_k\| \geq 3N - \|f\| \geq 2N > N$. Thus f is not in the interior of B_N and so B_N is nowhere dense and $\mathcal{Y} = \cup_{N=1}^\infty B_N$.

Example 3.14. This example shows that the assumption $\mathcal{X}$ is a Banach space can not be relaxed to $\mathcal{X}$ is simply a normed vector space in Theorem 3.11.

Let $\mathcal{Y}$ be an infinite dimensional Banach space. Let λ be a discontinuous linear functional, whose existence is the content of Proposition 2.14. As an exercise, show that the function $\|\cdot\|_* : \mathcal{Y} \to [0,\infty)$ given by

$$\|x\|_* = \|x\|_{\mathcal{Y}} + |\lambda(x)|$$

is a norm on $\mathcal{Y}$. Let $\mathcal{X}$ denote the normed space $\mathcal{Y}$ with this norm; that is $\mathcal{X} = (\mathcal{Y}, \|\cdot\|_*)$. Let $T : \mathcal{X} \to \mathcal{Y}$ denote the identity map (so bijective). Let G denote the unit ball in $\mathcal{X}$. We claim 0 is not in the interior of G as a subset of $\mathcal{Y}$. Indeed, since λ is not continuous, there is a sequence (x_n) of unit vectors in $\mathcal{Y}$ such that $|\lambda(x)| \geq n$. Consequently, given $r > 0$ and choosing n sufficiently large, $\|\frac{r}{2}x_n\|_{\mathcal{Y}} < r$, but $\|\frac{r}{2}x_n\|_* > 1$. Thus $B(0,r) \not\subseteq G$ proving the claim.

A consequence of the argument is that $\mathcal{X}$ is not a Banach space. To verify this fact directly, let $y \notin \ker\lambda$ be given. By Proposition prop:bdd-lf-iff-cns, there is a sequence (x_n) from $\ker\lambda$ that converges to x (in $\mathcal{Y}$). Since $\|x_n\|_* = \|x\|_{\mathcal{Y}}$, the sequence (x_n) is Cauchy in $\mathcal{X}$. However, since $\|x_n - y\|_* = \|x_n - y\|_{\mathcal{Y}} - \lambda(y)$, the sequence (x_n) does not converge to y in $\mathcal{X}$. Now suppose (x_n) converged to some $z \in \mathcal{X}$. Thus $z \in \mathcal{Y}$ and $\|x_n - z\|_* = \|x_n - z\|_{\mathcal{Y}} + \lambda(z)$ converges to 0 from which it follows that $z = y$ and the proof is complete.

This result depends on the axiom of choice. In this proof, choice is smuggle in through the appeal to Propostion 2.14, whose proof in turn depended on the existence of a Hamel basis, which in turn uses Zorn's Lemma (choice). □

END Monday 2025-02-03 - except had not discussed the Banach Isomorphism Theorem.

3.5. **The Closed Graph Theorem.** Recall that the Cartesian product $\mathcal{X} \times \mathcal{Y}$ of Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ with its default product topology from Subsection 1.2.7. In particular, the product topology on $\mathcal{X} \times \mathcal{Y}$ is the coarsest topology that makes both coordinate projections $\pi_{\mathcal{X}}$ and $\pi_{\mathcal{Y}}$ from $\mathcal{X} \times \mathcal{Y}$ to $\mathcal{X}$ and $\mathcal{Y}$ respectively continuous. This topology is the same as that determined by the norms in equation (2).

Definition 3.15. The *graph* of a linear map $T : \mathcal{X} \to \mathcal{Y}$ between normed vector spaces is the set

$$G(T) := \{(x, y) \in \mathcal{X} \times \mathcal{Y} : y = Tx\}.$$

Observe that since T is a linear map, $G(T)$ is a linear subspace of $\mathcal{X} \times \mathcal{Y}$. The transformation T is *closed* if $G(T)$ is a closed subset of $\mathcal{X} \times \mathcal{Y}$. □

It is an easy exercise to show that $G(T)$ is closed if and only if whenever (x_n, Tx_n) converges to (x, y), we have $y = Tx$. Problem 3.2 gives an example where $G(T)$ is closed, but T is not continuous. On the other hand, the next theorem says that if $\mathcal{X}, \mathcal{Y}$ are complete (Banach spaces), then $G(T)$ is closed if and only if T is continuous.

Theorem 3.16 (The Closed Graph Theorem)**.** *If $\mathcal{X}, \mathcal{Y}$ are Banach spaces and $T : \mathcal{X} \to \mathcal{Y}$ is closed, then T is bounded.*

Proof. We prove T closed implies T is bounded, leaving the easy converse as an exercise. With the norm $\|(x, y)\|_\infty = \max\{\|x\|, \|y\|\}$ the vector space $\mathcal{X} \times \mathcal{Y}$ is a Banach space with the product topology. The coordinate projections $\pi_\mathcal{X}, \pi_\mathcal{Y}$ are bounded with norm one. Let π_1, π_2 be the coordinate projections $\pi_\mathcal{X}, \pi_\mathcal{Y}$ restricted to $G(T)$; explicitly $\pi_1(x, Tx) = x$ and $\pi_2(x, Tx) = Tx$. Note that π_1 is a bijection between $G(T)$ and $\mathcal{X}$ and in particular $\pi_1^{-1}(x) = (x, Tx)$. By hypothesis $G(T)$ is a closed subset of a Banach space and hence a Banach space. Thus π_1 is a bounded linear bijection between Banach spaces and therefore, by Corollary 3.12, $\pi_1^{-1} : \mathcal{X} \to G(T)$ is bounded. Since π_2 is bounded, $\pi_2 \circ \pi_1^{-1} : \mathcal{X} \to \mathcal{Y}$ is continuous. To finish the proof, observe $\pi_2 \circ \pi_1^{-1}(x) = \pi_2(x, Tx) = Tx$. □

3.6. **Problems.**

Problem 3.1. Show that there exists a sequence of open, dense subsets $U_n \subseteq \mathbb{R}$ such that $m(\bigcap_{n=1}^\infty U_n) = 0$.

Problem 3.2. Consider the linear subspace $\mathcal{D} \subseteq c_0$ defined by

$$\mathcal{D} = \{f \in c_0 : \lim_{n\to\infty} |nf(n)| = 0\}$$

and the linear transformation $T : \mathcal{D} \to c_0$ defined by $(Tf)(n) = nf(n)$.

a) Prove T is closed, but not bounded. b) Prove T is bijective and $T^{-1} : c_0 \to \mathcal{D}$ is bounded (and surjective), but not open. c) What can be said of $\mathcal{D}$ as a subset of c_0?

Problem 3.3. Suppose $\mathcal{X}$ is a vector space equipped with two norms $\|\cdot\|_1, \|\cdot\|_2$ such that $\|\cdot\|_1 \le \|\cdot\|_2$. Prove that if $\mathcal{X}$ is complete in both norms, then the two norms are equivalent.

Problem 3.4. Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. Provisionally, say that a linear transformation $T : \mathcal{X} \to \mathcal{Y}$ is *weakly bounded* if $f \circ T \in \mathcal{X}^*$ whenever $f \in \mathcal{Y}^*$. Prove, if T is weakly bounded, then T is bounded.

Problem 3.5. Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. Suppose (T_n) is a sequence in $B(\mathcal{X}, \mathcal{Y})$ and $\lim_n T_n x$ exists for every $x \in \mathcal{X}$. Prove, if T is defined by $Tx = \lim_n T_n x$, then T is bounded.

Problem 3.6. Suppose that $\mathcal{X}$ is a vector space with a countably infinite basis. (That is, there is a linearly independent set $\{x_n\} \subseteq \mathcal{X}$ such that every vector $x \in \mathcal{X}$ is expressed uniquely as a *finite* linear combination of the x_n's.) Prove there is no norm on $\mathcal{X}$ under which it is complete. (Hint: consider the finite-dimensional subspaces $\mathcal{X}_n := \text{span}\{x_1, \dots x_n\}$.)

Problem 3.7. The Baire Category Theorem can be used to prove the existence of (very many!) continuous, nowhere differentiable functions on $[0, 1]$. To see this, let E_n denote the set of all functions $f \in C[0, 1]$ for which there exists $x_0 \in [0, 1]$ (which may depend on f) such that $|f(x) - f(x_0)| \leq n|x - x_0|$ for all $x \in [0, 1]$. Prove the sets E_n are nowhere dense in $C[0, 1]$; the Baire Category Theorem then shows that the set of nowhere differentiable functions is second category. (To see that E_n is nowhere dense, approximate an arbitrary continuous function f uniformly by piecewise linear functions g, whose pieces have slopes greater than $2n$ in absolute value. Any function sufficiently close to such a g will not lie in E_n.)

Problem 3.8. Let $L^2([0, 1])$ denote the Lebesgue measurable functions $f : [0, 1] \to \mathbb{C}$ such that $|f|^2$ is in $L^1([0, 1])$. It turns out, as we will see later, that $L^2([0, 1])$ is a linear manifold (subspace of the vector space $L^1([0, 1])$), though this fact is not needed for this problem.

Let $g_n : [0, 1] \to \mathbb{R}$ denote the function which takes the value n on $[0, \frac{1}{n^3}]$ and 0 elsewhere. Show,

(i) if $f \in L^2([0, 1])$, then $\lim_{n\to\infty} \int g_n f \, dm = 0$;
(ii) $L_n : L^1([0, 1]) \to \mathbb{C}$ defined by $L_n(f) = \int g_n f \, dm$ is bounded, and $\|L_g\| = n$;
(iii) conclude $L^2([0, 1])$ is of the first category in $L^1([0, 1])$.

Problem 3.9. A *Banach space of functions* on a set X is a vector subspace B of the space of complex-valued functions on X with a norm $\|\cdot\|$ making B a Banach space such that, for each $x \in X$, the mapping $E_x : B \to \mathbb{C}$ defined by $E_x(f) = f(x)$ is continuous (bounded) and if $f(x) = 0$ for all $x \in X$, then $f = 0$.

Suppose $g : X \to \mathbb{C}$. Show, if $gf \in B$ for each $f \in B$, then the linear map $M_g : B \to B$ defined by $M_g f = gf$ is bounded.

Problem 3.10. Suppose $\mathcal{X}$ is a Banach space and $\mathcal{M}$ and $\mathcal{N}$ are closed subspaces. Show, if for each $x \in \mathcal{X}$ there exist unique $m \in \mathcal{M}$ and $n \in \mathcal{N}$ such that

$$x = m + n,$$

then the mapping $P : \mathcal{X} \to \mathcal{M}$ defined by $Px = m$ is bounded.

Problem 3.11. Let $\mathcal{X}$ be a Banach space and $\mathcal{M} \subseteq \mathcal{X}$ a closed subspace. A linear transformation $P : \mathcal{X} \to \mathcal{M}$ is called a *bounded projection* if it is bounded and $P(m) = m$ for all $m \in \mathcal{M}$. Prove that if $\mathcal{M}$ is a closed subspace and there exists a bounded projection $P : \mathcal{X} \to \mathcal{M}$, then there exists a *closed* subspace $\mathcal{N} \subseteq \mathcal{X}$ such that $\mathcal{M} \cap \mathcal{N} = \{0\}$ and $\mathcal{X} = \mathcal{M} + \mathcal{N}$. Show also that in this case there exists a bounded projection $Q : \mathcal{X} \to \mathcal{N}$.

Remark: Given a closed subspace $\mathcal{M} \subseteq \mathcal{X}$, we say $\mathcal{M}$ is *(topologically) complemented* if there exists a *closed* subspace $\mathcal{N} \subseteq \mathcal{X}$ such that $\mathcal{M} \cap \mathcal{N} = \{0\}$ and $\mathcal{M} + \mathcal{N} = \mathcal{X}$. Taken together, the last two problems show that a closed subspace $\mathcal{M} \subseteq \mathcal{X}$ is complemented if and only if there is a bounded projection $P : \mathcal{X} \to \mathcal{M}$. Not every subspace of a Banach space is necessarily complemented, for example c_0 is not complemented in ℓ^∞, though this is nontrivial to prove.

Problem 3.12. Here, for definiteness we take the scalar field $\mathbb{R}$.

Let $T : \ell^\infty \to \ell^\infty$ denote the *backward shift* operator defined by $Tf(n) = f(n+1)$

A bounded linear functional $\lambda : \ell^\infty \to \mathbb{R}$ satisfying,

(i) if $f \in \ell^\infty$ and $(f(n))$ converges, then $\lambda(f) = \lim_{n\to\infty} f(n)$; and
(ii) $\lambda(Tf) = \lambda(f)$

is a *Banach Limit*.

Prove

(a) Banach limits exist.
(b) If λ is a Banach limit and $f \in \ell^\infty$, then

$$\liminf f(n) \leq \lambda(f) \leq \limsup f(n).$$

A sequence $f \in \ell^\infty$ for which $(f(n))$ does not converge, but $\lambda(f) = \mu(f)$ for all Banach limits λ and μ is *almost convergent*. Show that g defined by $g(n) = (-1)^n$ is almost convergent. (Suggestion: given a Banach limit λ, consider $\lambda(g + Tg)$.

Problem 3.13. Prove that $\mathbb{Q}$ is not a G_δ set.

4. L^p SPACES

Throughout this section, $(X, \mathscr{M}, \mu)$ is a measure space and $X \neq \varnothing$.

Definition 4.1. For $0 < p < \infty$, let $\mathcal{L}^p(\mu)$ denote the space of measurable functions $f : X \to \mathbb{F}$ that satisfy

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p} < \infty.$$

Lemma 4.2. *Suppose* $0 < p < \infty$. *If* $f, g \in \mathcal{L}^p(\mu)$ *and* $c \in \mathbb{F}$, *then* $\|f + g\|_p^p \leq 2^p \left(\|f\|_p^p + \|g\|_p^p \right)$ *and* $\|cf\|_p = |c| \, \|f\|_p$. *Hence* $\mathcal{L}^p(\mu)$ *is a vector space.*

Later we will show, for $1 \le p \le \infty$, that $\|\cdot\|_p$ is a semi-norm on $\mathcal{L}^p(\mu)$.

Sketch of proof. The equality $\|cf\|_p = |c|\,\|f\|_p$ is immediate. As a pointwise inequality, $|f+g| \le |f| + |g| \le 2\max\{|f|,|g|\}$. Hence,

$$|f+g|^p \le 2^p \max\{|f|,|g|\}^p = 2^p \max\{|f|^p,\, |g|^p\} \le 2^p(|f|^p + |g|^p),$$

from which the rest of the result follows. □

END Wednesday 2025-02-05

Definition 4.3. A measurable function $f : X \to \mathbb{F}$ is *essentially bounded* if there is a $t > 0$ such that

$$\mu(\{|f| > t\}) = 0$$

and let $\mathcal{L}^\infty(\mu)$ denote the set of essentially bounded functions on $(X, \mathscr{M}, \mu)$.

Define $\|\cdot\|_\infty : \mathcal{L}^\infty(\mu) \to [0,\infty)$ by

$$\|f\|_\infty = \inf\{t > 0 : \mu(\{|f| > t\}) = 0\}. \tag{15}$$

Proposition 4.4. *The set $\mathcal{L}^\infty(\mu)$ is a vector space. Further, the infimum in equation (15) is attained and $\|\cdot\|_\infty$ is a semi-norm on $\mathcal{L}^\infty(\mu)$.*

It is customary to write $\mathcal{L}^p$ instead of $\mathcal{L}^p(\mu)$ when the μ is understood (or generic).

Proof. It is evident that, if $c \in \mathbb{F}$ and $f \in \mathcal{L}^\infty$ then $cf \in \mathcal{L}^\infty$ and $\|cf\|_\infty = |c|\,\|f\|_\infty$. Now suppose $f, g \in \mathcal{L}^\infty$. Let $s, t > 0$ be given such that $s > \|f\|_\infty$ and $t > \|g\|_\infty$. By definition, the (measurable) sets $A = \{|f| > t\}$ and $B = \{|g| > s\}$ have measure 0. Let $C = \{|f+g| > s+t\}$. Hence $A \cup B$ has measure 0 and by the triangle inequality $A^c \cap B^c \subseteq C^c$. Thus $C \subseteq A \cup B$ and hence C has measure 0. Thus $f + g \in \mathcal{L}^\infty(\mu)$ and $\|f+g\| \le s+t$. It now follows that $\|f+g\| \le \|f\|_\infty + \|g\|_\infty$ and hence $\mathcal{L}^\infty$ is a vector space and $\|\cdot\|_\infty$ is a semi-norm on $\mathcal{L}^\infty$.

That the infimum is attained in equation (15) is left as an (easy) exercise based upon the fact that a countable union of sets of measure zero has measure zero. □

We record the following simple observation for later use - often without comment.

Lemma 4.5. *If $0 < p \le \infty$ and $f \in \mathcal{L}^p(\mu)$, then $\|f\|_p = 0$ if and only if $f = 0$ almost everyewhere.*

Proof. For $0 < p < \infty$, by assumption $g = |f|^p$ is unsigned $\int g = \|f\|_p^p$. Since $g = 0$ almost everywhere if and only if $\int g = 0$, the result follows. □

4.1. **Conjugate indices and the inequalities of Young, Holder and Minkowski.** We now restrict our attention to $1 \le p \le \infty$.

Definition 4.6. The *conjugate index* or *dual exponent* to $1 < p < \infty$ is the unique $1 < q < \infty$ satisfying

$$\frac{1}{p} + \frac{1}{q} = 1.$$

The dual index to $p = \infty$ is $q = 1$; and the dual index to $p = 1$ is $q = \infty$. □

Note that $(p-1)q = p$ and likewise $(q-1)p = q$. The significance of dual indices is apparent in the following result.

Lemma 4.7 (Young's inequality)**.** *If a, b are nonnegative numbers and $1 < p, q < \infty$ are dual indices, then*

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}$$

and equality holds if and only if $b^q = a^p$.

Proof. If a or b is 0 there is nothing to prove. So suppose $a, b > 0$. Define $\psi : \mathbb{R} \to \mathbb{R}$ by $\psi(t) = a^{p(1-t)} b^{qt}$. A bit of rearranging gives $\psi(t) = a^p \cdot \exp(ct)$, where $c = \log(b^q/a^p)$. The function ψ is infinitely differentiable and

$$\psi''(t) = c^2\, \psi(t) > 0.$$

Thus ψ is convex. In particular, (using the fact that $\frac{1}{p} + \frac{1}{q} = 1$)

$$\psi\left(\frac{1}{q}\right) = \psi\left(\frac{1}{p} \cdot 0 + \frac{1}{q} \cdot 1\right) \le \frac{1}{p}\psi(0) + \frac{1}{q}\psi(1) = \frac{a^p}{p} + \frac{b^q}{q}. \tag{16}$$

For the case of equality, note that $\psi(t)$ is strictly convex unless $c = 0$ ($a^p = b^q$), in which case ψ is constant. □

For an alternate geometric proof of Lemma 4.7, see Problem 2.9.

Theorem 4.8 (Hölder's inequality)**.** *Suppose $1 \le p \le \infty$ and q is the conjugate index to p. If $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$, then $fg \in \mathcal{L}^1$, and*

$$\|fg\|_1 \le \|f\|_p \|g\|_q. \tag{17}$$

Further, assuming $1 \le p < \infty$ and $f \in \mathcal{L}^p(\mu)$, if $\|f\|_p \neq 0$, then there exists a $g \in \mathcal{L}^q(\mu)$ such that

(i) $\|g\|_q = 1$;
(ii) $fg \ge 0$; *and*
(iii)

$$\|fg\|_1 = \int fg = \|f\|_p\, \|g\|_q = \|f\|_p. \tag{18}$$

If μ has the property that every set of positive measure contains a set of positive, but finite, measure and $f \in \mathcal{L}^\infty(\mu)$, then

$$\|f\|_\infty = \sup\{\|fg\|_1 : g \in \mathcal{L}^1(\mu),\ \|g\|_1 = 1\}.$$

Remark 4.9. For $f \in \mathcal{L}^1(\mu)$ of course equality holds in equation (17) with $g = 1$.

The assumption that $(X, \mathscr{M}, \mu)$ has the property that every (measurable) set of positive measure contains a set of finite measure is needed as the following example shows. For the measure space $(\{0\}, \{\varnothing, \{0\}\}, \mu)$, where $\mu(\varnothing) = 0$ and $\mu(\{0\}) = \infty$, we have $\mathcal{L}^1(\mu) = \{0\}$ and thus $\|fg\|_1 = 0$ for all $f \in \mathcal{L}^\infty(\mu)$ and $g \in \mathcal{L}^1(\mu)$. □

Proof of Theorem 4.8. The proof is easy in the cases $p = \infty$ or $p = 1$. Now suppose $1 < p < \infty$.

If $\|f\|_p = 0$, then $f = 0$ a.e. by Lemma 4.5. Hence $fg = 0$ a.e. and, by another application of Lemma 4.5, $\|fg\|_1 = 0$. Thus the inequality of equation (17) holds. By symmetry, the same is true for g. Hence we may assume $\|f\|_p \neq 0 \neq \|g\|_q$.

By homogeneity we may assume $\|f\|_p = \|g\|_q = 1$. We are to show

$$\int |fg|\, d\mu \leq 1.$$

Applying Lemma 4.7 gives

$$|f(x)g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q. \tag{19}$$

Integrating (19) with respect to μ and applying the normalizations on p, q, f, g gives the inequality of equation (17). Further, observe, in the case $1 < p, q < \infty$, that equality holds in Hölders in inequality if and only if equality holds a.e. μ in equation (19) if and only if $|f|^p = |g|^q$ a.e. μ by Lemma 4.7.

To prove the further portion of the theorem, suppose $1 < p < \infty$ and $f \in \mathcal{L}^p(\mu)$ satisfies $\|f\|_p = 1$. Let $g = |f|^p f^{-1}$ (interpreting g as 0 when f is 0). From $|g|^q = |f|^{(p-1)q} = |f|^p$ it follows that $g \in \mathcal{L}^q(\mu)$ and $\|g\|_q = 1$. Further, $fg = |f|^p$ so that $\|fg\|_1 = 1 = \|f\|_p$ and hence equation (18) holds. (Note that a small tweak to this argument also handles the case $p = 1$.)

For the last statement, suppose every subset S of X with $\mu(S) = \infty$ contains a set T for which $0 < \mu(T) < \infty$ and let $f \in \mathcal{L}^\infty(\mu)$ be given. Without loss of generality, $\|f\|_\infty = C > 0$. Given $0 < \rho < C$, the set $E = \{|f| > \rho\}$ has positive measure. Thus there is a set $F \subseteq E$ such that $0 < \mu(F) < \infty$. Let $g = \frac{1}{\mu(F)}\chi_F$, where χ_F is the characteristic function of F. Observe $g \in \mathcal{L}^1(\mu)$ and $\|g\|_1 = 1$. Moreover, $|fg| \geq \rho g$ and hence $\|fg\|_1 \geq \rho\|g\|_1$ and the result follows. □

END Friday 2025-02-07

Example 4.10. One can get a more intuitive feel for what Hölder's inequality says by examining it in the case of step functions. Let E, F be sets of finite, positive measure and put $f = 1_E, g = 1_F$. Then $\|fg\|_1 = \mu(E \cap F)$ and

$$\|f\|_p \|q\|_q = \mu(E)^{1/p} \mu(F)^{1/q},$$

so Hölder's inequality can be proved easily in this case using the relation $\frac{1}{p} + \frac{1}{q} = 1$ and the fact that $\mu(E \cap F) \le \min(\mu(E), \mu(F))$.

Corollary 4.11 (Minkowski's inequality)**.** *Let $(X, \mathscr{M}, \mu)$ be a measure space and suppose $1 \le p \le \infty$. If $f, g \in \mathcal{L}^p(\mu)$, then*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p.$$

Hence $\|\cdot\|_p$ is a semi-norm on $\mathcal{L}^p(\mu)$.

Proof. The result has already been established for $p = 1$ and $p = \infty$ so suppose $1 < p < \infty$ and let q denote the conjugate index to p. By Lemma 4.2, $f + g \in \mathcal{L}^p(\mu)$. The result is vacuous if $f + g = 0$ (almost everywhere); equivalenlty, $\|f + g\|_p = 0$. Now suppose $\|f + g\|_p \ne 0$. By Theorem 4.8, there is an $h \in \mathcal{L}^q(\mu)$ such that $\|h\|_q = 1$ and $\|(f + g)h\|_1 = \|f + g\|_p$. On the other hand, two more applications of Theorem 4.8 and the fact that $\|\cdot\|_1$ is a semi-norm gives,

$$\|f + g\|_p = \|(f + g)h\|_1 \le \|fh\|_1 + \|gh\|_1 \le \|f\|_p \|h\|_1 + \|g\|_p \|h\|_1 = \|f\|_p + \|g\|_p. \quad \square$$

4.2. **The Lebesuge spaces $L^p(\mu)$.** The proof of the following proposition, based on Lemma 4.5 is left to the gentle reader.

Proposition 4.12. *The set $\mathcal{N}(\nu) = \{f \in \mathcal{L}^p(\mu) : \|f\|_p = 0\}$ is a subspace of $\mathcal{L}^p(\mu)$ and the function $\|\cdot\|_p$ descends to a norm on the quotient space $\mathcal{L}^p(\mu)/\mathcal{N}(\mu)$.*

Definition 4.13. The normed vector space $(\mathcal{L}^p(\mu)/\mathcal{N}(\mu), \|\cdot\|_p)$ (for $1 \le p \le \infty$) is denoted $L^p(\mu)$ and is known as a *Lebesgue space*.

Suppose $1 \le p \le \infty$ and q is the conjugate index to p. Fix $g \in L^q(\mu)$. For $f \in L^p(\mu)$, Hölder's inequality (Theorem 4.8) implies $gf \in L^1(\mu)$ and moroever $\|fg\|_1 \le \|f\|_p \|g\|_q$. Thus, we obtain a bounded linear functional $L_g : L^p(\mu) \to \mathbb{F}$ of norm at most $\|g\|_q$ defined by

$$L_g(f) = gf.$$

Hence we obtain a bounded map (with norm at most one) $\Phi : L^q(\mu) \to L^p(\mu)^*$.

Proposition 4.14. *For $1 < p \le \infty$, the mapping $\Phi : L^q(\mu) \to L^p(\mu)^*$ defined by $\Phi(g) = L_g$ is isometric.*

When $p = 1$, if μ is σ-finite, then $\Phi : L^\infty(\mu) \to L^1(\mu)^*$ is isometric.

Remark 4.15. Returning to the example in Remark 4.9 where $\mathcal{L}^1(\mu) = \{0\}$, the map $\Phi : L^\infty(\mu) \to L^1(\mu)^*$ is not one-one. Since in this case $\mathcal{L}^1(\mu)^* = \{0\}$, but $\mathcal{L}^\infty(\mu) = \mathbb{F}$ isometrically, Φ is the zero map and, in particular, $\Phi(1) = 0$.

On the other hand, in the case $p = 1$ it suffices to assume that $(X, \mathscr{M}, \mu)$ has the property that every set S such that $\mu(S) = \infty$ contains a subset T with $0 < \mu(T) < \infty$.

Later we will see that the map Φ in Proposition 4.14 is an isometric isomorphism for $1 \le p < \infty$, with the proviso that μ is σ-finite in the case $p = 1$.

Problems 2.12 and 4.6 says that Φ need not be onto in the case that $p = \infty$.

Proof of Proposition 4.14. Let $g \in L^q(\mu)$ be given. If $\|g\|_q = 0$, then $L_g = 0$ so that the results holds, even when $p = 1$ without conditions on the measure space $(X, \mathscr{M}, \mu)$.

Now suppose $1 < p \le \infty$ and $\|g\|_q \neq 0$. As already observed, $\|\Psi(g)\| = \|L_g\| \le \|g\|_q$, for $g \in L^q(\mu)$. On the other hand, since $1 \le q < \infty$, for $0 \neq g \in L^q(\mu)$, the moreover portion of Hölder's Inequality (Theorem 4.8) gives a function $f \in L^p(\mu)$ such that $\|f\|_p = 1$ and

$$L_g(f) = \int fg = \|g\|_q.$$

Hence $\|L_g\| \ge \|g\|_q$ and thus $\|L_g\| = \|g\|_q$.

In the case $p = 1$ and μ satisfies the hypothesis of the additional hypotheses given, Hölder's inequality implies that, for each $0 \le \rho < \|g\|_\infty$, there is an $f \in L^1(\mu)$ such that $\|f\|_1 = 1$ and $|L_g(f)| = \|fg\|_1 > \rho$. Thus $\|L_g\| \ge \|g\|_\infty$ and consequently $\|L_g\| = \|g\|_\infty$. ☐

Proposition 4.16. *For $1 \le p \le \infty$, the normed vector space $L^p(\mu)$ is a Banach space.*

Proposition 4.16 is a near immediate consequence of the following lemma.

Lemma 4.17. *Suppose $1 \le p \le \infty$ and $(f_n)_{n=1}^\infty$ is a sequence from $\mathcal{L}^p(\mu)$. If for each $\epsilon > 0$ there is an N such that if $m, n \ge N$, then $\|f_n - f_m\|_p < \epsilon$, then there is an $f \in \mathcal{L}^p(\mu)$ such that*

(a) the sequence $(\|f_n - f\|_p)_n$ converges to 0;
(b) there is subsequence (f_{n_k}) of (f_n) that converges to f pointwise almost everywhere.

END Monday 2025-02-10

Sketch of proof. The proof for the case $1 \le p < \infty$ is very much like the case $p = 1$ that has already been established and is just sketched here.

There is a subsequence $(g_k)_{k=1}^\infty$ of (f_n) such that $\|g_{k+1} - g_k\| < 2^{-k}$ for $k \ge 1$. Setting $g_0 = 0$, the series $\sum_{k=0}^\infty \|g_{k+1} - g_k\|_p$ converges. (The subsequence (g_k) is super-Cauchy.) Let

$$h_m = \sum_{k=0}^{m} |g_{k+1} - g_k|$$

and let h denote the pointwise limit (in $[0,\infty]$) of the non-negative increasing sequence (h_m). By the Monotonne Convergence Theorem,

$$\int h^p = \lim \int h_m^p.$$

Thus,

$$\|h\|_p = \lim \|h_m\|_p.s \tag{20}$$

The inequality,

$$\|h_m\|_p \le \sum_{k=0}^{m} \|g_{k+1} - g_k\|_p \le \left[\sum_{k=0}^{\infty} \|g_{k+1} - g_k\|_p\right]^p < \infty$$

and equation (20) implies $h \in \mathcal{L}^p$. Thus

$$h = \sum_{k=0}^{\infty} |g_{k+1} - g_k|$$

is finite almost everywhere and hence,

$$\sum_{k=0}^{m} (g_{k+1} - g_k) = g_{m+1}$$

also converges almost everywhere to a measurable function f. That is, the sequence (g_k) converges pointwise to f. Further, since $|f| \le h$ and $h \in \mathcal{L}^p$, it follows that $f \in \mathcal{L}^p$.

By construction, for m fixed, if $n > m$, then $\|g_n - g_m\|_p < 2^{1-m}$ and $(|g_n - g_m|^p)_n$ converges pointwise almost everywhere to $|f - g_m|^p$. Thus, by Fatou,

$$\|f - g_m\|_p^p = \int |f - g_m|^p \le \liminf \int |g_n - g_m|^p = \liminf \|g_n - g_m\|_p^p < 2^{1-m}.$$

Thus the sequence $(\|f - g_m\|_p)_m$ converges to 0.

A standard fact that, in a metric space X, if (x_n) is Cauchy and if there is an $x \in X$ and a subsequence (x_{n_k}) of (x_n) that converges to x, then (x_n) converges to x. Thus, from what has been proved, (f_n) converges to f in $\mathcal{L}^p$ and the subsequence (g_k) of (f_n) converges to f pointwise, completing the proof for $1 \le p < \infty$.

The case $p = \infty$ follows from the fact that, for $g \in \mathcal{L}^\infty(\mu)$, the set $\{|g| > \|g\|_\infty\}$ has measure 0 (Proposition 4.4) so that it can be assumed that $|g|$ is bounded by $\|g\|_\infty$. From here the proof is very much like the proof of completeness of the space $F_b(X, \mathbb{F})$ of bounded functions on a set X with the supremum norm. (See Proposition 1.16) In particular, a Cauchy sequence (f_n) converges pointwise almost everywhere (no subsequence is needed). The details are left to the reader. □

Corollary 4.18. *If (f_n) is a sequence from $\mathcal{L}^p(\mu)$ $(1 \le p \le \infty)$ that converges to f in $\mathcal{L}^p(\mu)$, then there is a subsequence (f_{n_k}) of (f_n) that converges to f pointwise almost everywhere.*

Proof. The sequence (f_n) (viewed as representative of their respective equivalence classes) satisfies the hypotheses of Lemma 4.17. Hence there is a $g \in L^p(\mu)$ and subsequence (g_k) of (f_n) that converges to g both in $L^p(\mu)$ and pointwise almost everywere. By uniqueness of limits, $g = f$ as elements of $L^p(\mu)$; that is, almost everywhere. Thus (g_k) converge to f pointwise almost everywhere. □

Corollary 4.19. *Suppose $1 \leq p \leq \infty$ and (f_n) is a sequence from $L^p(\mu)$. If (f_n) converges to f in $L^p(\mu)$ and to g pointwise a.e., then $f = g$ a.e.; that is, the pointwise limit and $L^p(\mu)$ limit are the same (almost everywhere).*

Example 4.20. [The typewriter sequence] Define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \chi_{(\frac{n-2^k}{2^k}, \frac{n+1-2^k}{2^k}]}, \quad \text{for } 2^k \leq n < 2^{k+1},$$

(here χ is the indicator function) viewed as functions in $L^1(m)$ for Lebesgue measure m on $[0,1]$. The sequence (f_n) converges to 0 in $L^1(m)$, but does not converge pointwise anywhere. On the other hand, the subsequence $(g_k = f_{2^k})$ converges to 0 pointwise (everywhere).

Example 4.21. Let $(X, \mathscr{M}, \mu)$ denote a measure space and suppose $h : X \to \mathbb{F}$ is a measurable function. Given $1 \leq p, r \leq \infty$, if $hf \in L^r$ for each $f \in L^p$, then the linear mapping $M_h : L^p \to L^r$ is bounded. As an example of what more can be said, if $\mu(X) < \infty$ and $p = 2 = r$, then $h \in L^\infty$.

Use the Closed Graph Theorem as follows. Suppose $(f_n, M_h f_n)$ is a sequence that converges to (f, g) in $\mathcal{L}^p \times \mathcal{L}^r$. Apply Corollary 4.18 to (f_n) and f to obtain a subsequence (g_k) of (f_n) converging to f pointwise a.e. and of course in $\mathcal{L}^p$. Apply Corollary 4.18 to (hg_k) and g to deduce $g = hf$. Now use Closed Graph.

For the bit about L^2, make an argument like the one at the end of the proof of Hölder's inequality.

4.3. **Problems.**

Problem 4.1. Suppose $f : [0, A] \to [0, \infty)$ is differentiable, strictly increasing and $f(0) = 0$. Prove, for each $0 < a \leq A$, that

$$\int_0^x f + \int_0^{f(x)} f^{-1} = xf(x).$$

[Suggestion: Differentiate $g(x) = \int_0^x f + \int_0^{f(x)} f^{-1} - xf(x)$.] Deduce Young's inequality.

Problem 4.2. [Truncation of L^p functions] Suppose f is an unsigned function in $L^p(\mu)$, $1 < p < \infty$. For $t > 0$ let

$$E_t = \{x : f(x) > t\}.$$

Show:

(a) For each real number $t > 0$, the *horizontal truncation* $1_{E_t} f$ belongs to L^q for all $1 \le q \le p$.

(b) For each real number $t > 0$, the *vertical truncation* $f_t := \min(f, t)$ belongs to L^q for all $p \le q \le \infty$.

(c) Every $f \in L^p$, $1 < p < \infty$, can be decomposed as $f = g + h$ where $g \in L^1$ and $h \in L^\infty$.

Problem 4.3. Suppose $f \in L^{p_0} \cap L^\infty$ for some $p_0 < \infty$. Prove $f \in L^p$ for all $p_0 \le p \le \infty$, and $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$.

Problem 4.4. Prove $f_n \to f$ in the L^∞ norm if and only if $f_n \to f$ essentially uniformly, and that L^∞ is complete.

Problem 4.5. Show that $L^p(\mathbb{R}) \not\subseteq L^q(\mathbb{R})$ for any pair p, q.

Problem 4.6. Consider $L^\infty(\mathbb{R})$.

a) Show that $\mathcal{M} := C_0(\mathbb{R})$ is a closed subspace of $L^\infty(\mathbb{R})$ (more precisely, that the set of L^∞ functions that are a.e. equal to a C_0 function is closed in L^∞). Prove there is a bounded linear functional $\lambda : L(\mathbb{R})^\infty \to \mathbb{F}$ such that $\lambda|_{\mathcal{M}} = 0$ and $\lambda(1_\mathbb{R}) = 1$.
b) Prove there is no function $g \in L^1(\mathbb{R})$ such that $\lambda(f) = \int_\mathbb{R} fg\,dm$ for all $f \in L^\infty$. (Hint: look at the restriction of λ to $C_0(\mathbb{R})$.)

5. Hilbert Space

5.1. Inner product spaces.

Definition 5.1. Let V denote a vector space over $\mathbb{C}$. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ is a *inner product* if, for all $f, g, h \in V$ and $c \in \mathbb{C}$,

(a) $\langle f, f \rangle \ge 0$;
(b) $\langle f, f \rangle = 0$ if and only if $f = 0$;
(c) $\langle f + cg, h \rangle = \langle f, h \rangle + c \langle g, h \rangle$;
(d) $\langle g, f \rangle = \overline{\langle f, g \rangle}$.

END Wednesday 2025-02-12

Proposition 5.2. *An inner product on a vector space V satisfies the the* Cauchy–Schwarz inequality*,*

$$|\langle f, g \rangle|^2 \le \langle f, f \rangle \langle g, g \rangle.$$

Equality holds if and only if f and g are linearly dependent.

The function $\|\cdot\| : V \to \mathbb{C}$ defined by $\|f\| = \sqrt{\langle f, f \rangle}$ is a norm on V and, with this notation, the CS inequality becomes

$$|\langle f, g \rangle| \le \|f\| \, \|g\|.$$

Further, $\|f+g\| = \|f\| + \|g\|$ *if and only if either* $f = 0$ *or there is a* $t \geq 0$ *such that* $g = tf$*.*

Remark 5.3. Given an inner product space $V = (V, \langle\cdot,\cdot\rangle)$, we endow it with the norm - and hence metric - arising from the inner product.

Lemma 5.4 (Joint continuity of the inner product)**.** *Let* H *be an inner product space equipped with its norm topology. If* (x_n) *converges to* x *and* (y_n) *converges to* y *in* H*, then* $(\langle x_n, y_n\rangle)$ *converges to* $\langle x, y\rangle$*.*

Proof. By Cauchy-Schwarz,

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \leq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \to 0,$$

since $\|x_n - x\|, \|y_n - y\| \to 0$ and the sequence $\|x_n\|$ is bounded. □

Definition 5.5. A *Hilbert space* H over $\mathbb{F}$ is an inner product space over $\mathbb{F}$ that is complete in the metric $d(x,y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle}$. (Here, as usual, $\mathbb{F}$ is either $\mathbb{C}$ or $\mathbb{R}$.)

We continue to use the notation $M \leq H$ to indicate M is a (closed) subspace of H from Definition 1.26.

Example 5.6 ($\mathbb{F}^n$)**.** It is easy to check that the standard scalar product on $\mathbb{R}^n$ is an inner product; it is defined as usual by

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j \tag{21}$$

where we have written $x = (x_1, \dots x_n)$; $y = (y_1, \dots y_n)$. Similarly, the standard inner product of vectors $z = (z_1, \dots z_n)$, $w = (w_1, \dots w_n)$ in $\mathbb{C}^n$ is given by

$$\langle z, w\rangle = \sum_{j=1}^{n} z_j \overline{w_j}. \tag{22}$$

(Note that it is necessary to take complex conjugates of the w's to obtain positive definiteness.)

It is straightforward to check that equations (21) and (22) define inner products on $\mathbb{R}^n$ and $\mathbb{C}^n$ respectively that induce the Euclidean norm. Since these Euclidean spaces are complete, they are Hilbert spaces.

Example 5.7 ($L^2(\mu)$)**.** Let $(X, \mathscr{M}, \mu)$ be a measure space. Given $f, g \in L^2(\mu)$, by Hölder's inequality (Theorem 4.8), the function $f\overline{g} \in L^1(\mu)$ and $\|f\overline{g}\| \leq \|f\|_2 \|g\|_2$. From here it is a simple exercise to verify that the Banach space $L^2(\mu)$ is the inner product space with the inner product,

$$\langle f, g\rangle = \int_X f\overline{g}\, d\mu. \tag{23}$$

That is, equation (23) is an inner product and the norm on $L^2(\mu)$ is the norm derived from this inner product.

Example 5.8 ($\ell^2(\mathbb{N})$)**.** Let

$$\ell^2(\mathbb{N}) = \{(a_1, a_2, \dots a_n, \dots) \mid a_n \in \mathbb{F},\ \sum_{j=1}^{\infty} |a_n|^2 < \infty\}.$$

This space is $L^2(c)$ for the measure space $(\mathbb{N}, P(\mathbb{N}), c)$, where c is counting measure. In particular, $\ell^2(\mathbb{N})$ is a Hilbert space with the inner product,

$$\sum_{n=1}^{\infty} a_n \overline{b_n} \tag{24}$$

for sequences $a = (a_1, a_2, \dots)$ and $b = (b_1, b_2, \dots)$ in ℓ^2.

Note too, that example 5.6, is the special case of $L^2(\nu)$ for ν equal to counting measure on $P(\{1, 2, \dots, n\})$.

5.2. **Orthogonality.** In this section we show that many of the basic features of the Euclidean geometry of $\mathbb{F}^n$ extend naturally to the setting of an inner product space.

Definition 5.9. Let H be an inner product space.

(i) Two vectors $x, y \in H$ are *orthogonal* if $\langle x, y\rangle = 0$, written $x \perp y$.
(ii) Two subsets A, B of H are *orthogonal* if $x \perp y$ for all $x \in A$ and $y \in B$, written $A \perp B$.
(iii) A subset A of H is *orthogonal* if $x \perp y$ for each $x, y \in A$ with $x \neq y$ and is *orthonormal* if also $\langle x, x\rangle = 1$ for all $x \in A$.
(iv) The *orthogonal complement* of a subset E of H is

$$E^\perp = \{x \in H : \langle x, e\rangle = 0 \text{ for all } e \in E\}.$$

The proof of the following lemma is an easy exercise. Indeed, the first item follows immediately from Lemma 5.4 and the second from the positive definiteness of a norm.

Lemma 5.10. *If E is a subset of an inner product space H, then*

(i) $E^\perp$ is a closed subspace of H;
(ii) $E \cap E^\perp \subseteq \{0\}$; and
(iii) $E \subseteq (E^\perp)^\perp = E^{\perp\perp}$.

Theorem 5.11 (The Pythagorean Theorem)**.** *If H is an inner product space and $f_1, \dots f_n$ are mutually orthogonal vectors in H, then*

$$\|f_1 + \cdots + f_n\|^2 = \|f_1\|^2 + \cdots + \|f_n\|^2.$$

Proof. When $n = 2$, we have

$$\begin{aligned}\|f_1 + f_2\|^2 &= \|f_1\|^2 + \langle f_1, f_2\rangle + \langle f_2, f_1\rangle + \|f_2\|^2 \\ &= \|f_1\|^2 + \|f_2\|^2.\end{aligned}$$

The general case follows by induction. □

Suppose V is a vector space over $\mathbb{F}$. A function $[\cdot,\cdot] : V \times V \to \mathbb{F}$ satisfying the axioms of items c and d is *bilinear form* in the case $\mathbb{F} = \mathbb{R}$ and a *sesquilinear form* when $\mathbb{F} = \mathbb{C}$. If it also satisfies item a, then it is *positive semi-defininite*.

Theorem 5.12 (The Parallelogram Law)**.** *If $[\cdot,\cdot]$ is a bilinear (resp. sesquilinear) form on a vector space over $\mathbb{R}$ (resp. $\mathbb{C}$) and $f, h \in V$, then*

$$[f+g, f+g] + [f-g, f-g] = 2\left([f,f] + [g,g]\right). \tag{25}$$

In particular, if H is an inner product space, then

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|^2 + \|g\|^2). \tag{26}$$

Proof. By linearity (resp. sesquilinearity),

$$[f \pm g, f \pm g] = [f,f] \pm [f,g] \pm [g,f] \pm [g,g] \tag{27}$$

Adding these two equations together gives the identity of equation (25).

In the case of Hilbert space, equation (26) follows from equation (25) by the definition of the norm coming from the inner product. □

Subtracting, instead of adding, in the proof of the Parallelogram Law gives the polarization identity

$$2\left([f,g] + [g,f]\right) = [f+g, f+g] - [f-g, f-g].$$

Theorem 5.13 (The Polarization identity)**.** *If $[\cdot,\cdot]$ is a bilinear form on a vector space over $\mathbb{R}$and $f, h \in V$, then*

$$4[f,g] = [f+g, f+g] - [f-g, f-g]. \tag{28}$$

In particular, if H is an inner product space over $\mathbb{R}$, then

$$\langle f, g\rangle = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2\right). \tag{29}$$

If $[\cdot,\cdot]$ is a sesquilinear form on a vector space over $\mathbb{C}$and $f, h \in V$, then

$$4[f,g] = \sum_{k=0}^{3} i^k [f + i^k g, f + i^k g]. \tag{30}$$

If H is a complex Hilbert space, then

$$4\langle f, g\rangle = \sum_{k=0}^{3} i^k \langle f + i^k g, f + i^k g\rangle. \tag{31}$$

Remark: Note that, in a Hilbert space, the polarization identity says that the inner product is determined by the norm.

An elementary (but tricky) theorem of von Neumann says, in the real case, that if H is any vector space equipped with a norm $\|\cdot\|$ such that the parallelogram law (26) holds for all $f, g \in H$, then H is an inner product space with inner product given by formula (29) in the case of real scalars and formula (31) in the case of complex scalars. (The proof is simply to *define* the inner product by equation (29) or (31), and check that it is indeed an inner product.)

5.3. **Best approximation in Hilbert space.**

Definition 5.14. A subset K of a vector space V is *convex* if whenever $a, b \in K$ and $0 \le s, t$ sum to 1, it follows that $sa + tb \in K$ as well. (Geometrically, this means that when a, b lie in K, so does the line segment joining them.)

A normed vector space $\mathcal{X}$ is *strictly convex* if $x, y \in \mathcal{X}$ and $\|x + y\| = \|x\| + \|y\|$, then either $x = 0$ or there is a $t \ge 0$ such that $y = tx$.

Example 5.15. Subspaces and balls ($B(x, r)$) in a normed vector spaces are convex. The closure and interior of a convex set are convex.

Remark 5.16. Hilbert spaces are strictly convex. The Lebesgue spaces L^p are convex for $1 < p < \infty$, but not for $p = 1, \infty$.

That a normed vector space is strictly convex if and only if $x \neq y$ and $\|x\| = \|y\| = 1$, then $\|\frac{1}{2}(x + y)\| < 1$ offers an explanation for the terminology.

END Friday 2025-02-14 (though we had not finished with the remark immediately above)

Proposition 5.17. *Suppose $\mathcal{X}$ is a strictly convex normed vector space. If $K \subseteq \mathcal{X}$ is convex, $h \in \mathcal{X}$ and there exists a $y, z \in K$ such that*

$$\|h - y\| = \operatorname{dist}(h, K) = \inf\{\|h - k\| : k \in K\} = \|h - z\|,$$

then $z = y$.

Proof. Let $d = \operatorname{dist}(h, K)$. By convexity, $k = \frac{y+z}{2} \in K$ and by the triangle inequality,

$$d \le \|h - k\| = \|\frac{1}{2}(h - y) + \frac{1}{2}(h - z)\| \le \frac{1}{2}\left[\|h - y\| + \|h - z\|\right] = d.$$

Hence equality holds in the triangle inequality. Without loss of generality, $h - y \neq 0$ and, by strict convexity, there is a $t \ge 0$ such that $h - y = t(h - z)$. Since $\|h - y\| = d = \|h - z\|$, it follows that $t = 1$ and therefore $y = z$. ☐

Theorem 5.18. *Suppose H is a Hilbert space. If $\varnothing \neq K \subseteq H$ is a closed, convex, nonempty set, and $h \in H$, then there exists a unique vector $k_0 \in K$ such that*

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

Proof. Uniqueness follows from Propositions 5.2 and 5.17.

Let $d = \operatorname{dist}(h, K) = \inf_{k\in K} \|h-k\|$. First observe, if $x, y \in K$, then, by convexity, so is $v = \frac{x+y}{2}$ and in particular, $\|h-v\|^2 \geq d^2$. Hence, by the parallelogram law, applied to $f = \frac{x-h}{2}$ and $g = \frac{y-h}{2}$,

$$\begin{aligned} \left\|\frac{x-y}{2}\right\|^2 =& \frac{1}{2}\left(\|x-h\|^2 + \|y-h\|^2\right) - \left\|\frac{x+y}{2} - h\right\|^2 \\ \leq& \frac{1}{2}\left(\|x-h\|^2 + \|y-h\|^2\right) - d^2. \end{aligned} \tag{32}$$

There exists a sequence (k_n) in K so that $(\|k_n - h\|)$ converges to d. Given $\epsilon > 0$ choose N such that for all $n \geq N$, $\|k_n - h\|^2 < d^2 + \epsilon^2$. By (32), if $m, n \geq N$ then

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

Consequently $\|k_m - k_n\| < \epsilon$ for $m, n \geq N$ and (k_n) is a Cauchy sequence. Since H is complete, (k_n) converges to a limit $k \in H$, and since K is closed, $k \in K$. Since $(k_n - h)$ converges to $(k-h)$ and $\|k_n - h\|$ converges to d it follows, by continuity of the norm, that $\|k-h\| = d$. □

The most important application of the preceding approximation theorem is in the case when $K = M$ is a (closed) subspace of the Hilbert space H. What is significant is that in the case of a subspace, the minimizer k has an elegant geometric description, namely, it is obtained by "dropping a perpendicular" from h to M. This geometric interpretation is the content of the next theorem, whose statement uses Theorem 5.18. Recall $M \leq H$ to means that M is a (closed) subspace of H.

Theorem 5.19. *Suppose H is a Hilbert space, $M \leq H$, and $h \in H$. If f_0 is the unique element of M such that $\|h - f_0\| = \operatorname{dist}(h, M)$, then $(h - f_0) \perp M$. Conversely, if $f_0 \in M$ and $(h - f_0) \perp M$, then $\|h - f_0\| = \operatorname{dist}(h, M)$.*

Proof. Let $f_0 \in M$ with $\|h - f_0\| = \operatorname{dist}(h, M)$ be given. Given $f \in M$, for $t \in \mathbb{R}$, let $\lambda = t\langle h - f_0, f\rangle$. Since $f_0 + \lambda f \in M$,

$$\begin{aligned} 0 \leq& \|h - (f_0 + \lambda f)\|^2 - \|h - f_0\|^2 \\ =& \|(h - f_0) + \lambda f\|^2 - \|h - f_0\|^2 \\ =& -2 \operatorname{real} \overline{\lambda} \langle h - f_0, f\rangle + |\lambda|^2 \|f\|^2 \\ =& \left[-2t + t^2 \|f\|^2\right] \left|\langle h - f_0, f\rangle\right|^2 \end{aligned}$$

for all t. Thus $|\langle h - f_0, f\rangle| = 0$.

Conversely, suppose $f_0 \in M$ and $(h - f_0) \perp M$. In particular, we have $(h - f_0) \perp (f_0 - f)$ for all $f \in M$. Therefore, by Theorem 5.11, for all $f \in M$

$$\begin{aligned}\|h - f\|^2 &= \|(h - f_0) + (f_0 - f)\|^2 \\ &= \|h - f_0\|^2 + \|f_0 - f\|^2 \geq \|h - f_0\|^2.\end{aligned}$$

Thus $\|h - f_0\| = \operatorname{dist}(h, M)$. □

Corollary 5.20. *If H is a Hilbert space and $M \leq H$, then $(M^\perp)^\perp = M$.*

Proof. By Lemma 5.10, $M \subseteq (M^\perp)^\perp$. Now suppose that $x \in (M^\perp)^\perp$. By Theorem 5.19 applied to x and M, there exists $m \in M$ such that $x - m \in M^\perp$. On the other hand, both x and m are in $(M^\perp)^\perp$ and thus by Lemma 5.10, $x - m \in (M^\perp)^\perp$. Hence $x - m = 0$ by Lemma 5.10, and $x \in M$. □

If E is a subset of the Banach space X, and $\mathscr{E}$ is the collection of all closed subspaces $\mathcal{N}$ of X such that $E \subseteq \mathcal{N}$, then

$$\mathcal{M} = \cap_{\mathcal{N} \in \mathscr{E}} \mathcal{N}$$

is the *smallest closed subspace containing* E.

Corollary 5.21. *If E is a subset of H, then $(E^\perp)^\perp$ is equal to the smallest closed subspace of H containing E. In particular, if E is a linear manifold (vector subspace) in H, then $\overline{E} = (E^\perp)^\perp$.*

Proof. The proof uses Lemma 5.10 freely. In particular, $E \subseteq (E^\perp)^\perp$ and $(E^\perp)^\perp$ is a closed subspace. If M is a closed subspace containing E, then $E^\perp \supseteq M^\perp$ and hence $(E^\perp)^\perp \subseteq (M^\perp)^\perp = M$ by Corollary 5.20.

For the last statement, from the fact that $\overline{E}$ and $E^{\perp\perp}$ are both the smallest closed subspace containing the linear maniforld E. See Corollary 5.20. □

END Monday 2025-02-17

Corollary 5.22. *A vector subspace E of a Hilbert space H is dense in H if and only if $E^\perp = \{0\}$.*

Proposition 5.23. *Suppose $M, N \leq H$. If M and N are orthogonal, then $M + N$ is closed. In particular, $M + N$ is again a subspace of H.*

Proof. It suffices to prove that $M + N$ is complete. Accordingly suppose $(m_k + n_k)$ is a Cauchy sequence from $M + N$. From orthogonality, for $k, \ell \in \mathbb{N}$,

$$\|m_k - m_\ell\|^2 + \|n_k - n_\ell\|^2 = \|(m_k + n_k) - (m_\ell + n_\ell)\|^2$$

and hence (m_k) and (n_k) are both Cauchy. Since H is complete and M, N are closed, M and N are each complete. Thus (m_k) converges to some $m \in M$ and (n_k) converges to some $n \in N$ and thus $(m_k + n_k)$ converges to $m + n \in M + N$. □

Definition 5.24. Given subspaces $M, N \le H$ of a Hilbert space H, the notation $M \oplus N$ is used for $M + N$ in the case M and N are closed subspaces and $M \perp N$ and is called the *orthogonal direct sum*. Hence, $M \oplus N$ indicates that M, N are orthogonal closed subspaces of H.

The following corollary should be compared with Problem 3.10.

Corollary 5.25. *If $M \le H$, then $H = M \oplus M^{\perp}$.*

Proof. Given $x \in H$, there exists $m \in M$ such that $x - m \in M^{\perp}$ by Theorem 5.19. Hence $x = m + (x - m) \in M \oplus M^{\perp}$. ☐

Example 5.26. In a Banach space, a *best* approximation to a subspace need not exist as the following example illustrates. Consider the real Banach space $C([0,1])$, the subspace $U = \ker \lambda_1 \cap \ker \lambda_2$ where λ_j are the linear functionals on $C([0,1])$ defined by $\lambda_1(f) = \int f$ and $\lambda_2(f) = f(1)$. Since these linear functionals are bounded with norm 1, the linear manifold U is closed (so a subspace). Let $f = 1 - x$ and observe, for $g \in U$, that $(f - g)(1) = 0$ and $\int (f - g) = \frac{1}{2}$. Thus the average of $f - g$ is $\frac{1}{2}$ but $(f - g)(1) < \frac{1}{2}$. Consequently, there is a point $p \in [0,1]$ such that $(f - g)(p) > \frac{1}{2}$ and we conclude that there does not exists a $g \in U$ such that $\|g - f\| = \frac{1}{2}$.

Given $0 < \epsilon < \frac{1}{2}$, choose $0 < \delta = \frac{\frac{1}{2} - \epsilon + 2\epsilon^2}{1 + 2\epsilon} < 1$ Let $\gamma = \frac{1}{2} - \epsilon - \delta$ and let $g = g_\epsilon$ denote the piecewise linear function that takes values $\frac{1}{2} - \epsilon - x$ for $0 \le x \le \delta$, and then connects the points (δ, γ) to $(1, 0)$ (draw the picture). By construction, $g \in C([0,1])$ and $g(1) = 0$. Further, δ was chosen to insure that $\int g = 0$. Thus $g \in U$ and $\|f - g\|_\infty = \frac{1}{2} + \epsilon$. Hence $\operatorname{dist}(f, U) = \frac{1}{2}$ but there does not exist a $g \in U$ where this distance is achieved.

Example 5.27. This example show that in a Banach space, there can be more than one closest point from a point to a subsapce.

Consider the real Banach space $(\mathbb{R}^2, \|\cdot\|_\infty)$ (thus $\|(x_1, x_2)\|_\infty = \max\{|x_1|, |x_2|\}$. Let $M = \{(x_1, 0) : x_1 \in \mathbb{R}\} \subseteq \mathbb{R}^2$ and note M is a subspace of $\mathbb{R}^2$. Let $y = (0, 1)$ and observe $\operatorname{dist}(y, M) = 1$ and this distance is attained for each $(x, 0) \in M$ with $|x| \le 1$.

END Wednesday 2025-02-19 - we also discussed Corollary 4.19 and example 4.20.

5.4. **The Riesz Representation Theorem and Hilbert space adjoint operators.** In this section we investigate the dual H^* of a Hilbert space H. One way to construct bounded linear functionals on Hilbert space is as follows. Given a vector $g \in H$ define,

$$L_g(h) = \langle h, g \rangle.$$

Indeed, linearity of L is just the linearity of the inner product in the first entry, and the boundedness of L follows from the Cauchy-Schwarz inequality,

$$|L_g(h)| = |\langle h, g \rangle| \le \|g\| \|h\|.$$

So $\|L_g\| \le \|g\|$, but in fact it is easy to see that $\|L_g\| = \|g\|$; just apply L_g to the unit vector $g/\|g\|$ (assuming $g \neq 0$). Hence, $L : H \to H^*$ defined by $g \mapsto L_g$ is a *conjugate linear* isometry (thus linear in the case of real scalars).

In fact, it is clear from linear algebra that every linear functional on $\mathbb{F}^n$ takes the form L_g. More generally, every *bounded* linear functional on a Hilbert space has the form just described.

Theorem 5.28 (The Riesz RepresentationTheorem)**.** *If H is a Hilbert space and $\lambda : H \to \mathbb{F}$ is a bounded linear functional, then there exists a unique vector $g \in H$ such that $\lambda = L_g$. Thus the conjugate linear mapping L is isometric and onto.*

Proof. It has already been established that L is isometric and in particular one-one. Thus it only remains to show L is onto. Accordingly, let $\lambda \in H^*$ be given. If $\lambda = 0$, then $\lambda = L_0$. So, assume $\lambda \neq 0$. Since λ is continuous, by Proposition 1.32 $\ker \lambda = \lambda^{-1}(\{0\})$ is a *proper, closed* subspace of H. Thus, by Theorem 5.19 (or Corollary 5.25) there exists a nonzero vector $f \in (\ker \lambda)^{\perp}$ and by rescaling we may assume $\lambda(f) = 1$.

Given $h \in H$, observe

$$\lambda(h - \lambda(h)f) = \lambda(h) - \lambda(h)\lambda(f) = 0.$$

Thus $h - \lambda(h)f \in \ker \lambda$ and consequently,

$$\begin{aligned} 0 &= \langle h - \lambda(h)f, f\rangle \\ &= \langle h, f\rangle - \lambda(h)\langle f, f\rangle. \end{aligned}$$

Thus $\lambda = L_g$, where $g = \frac{f}{\|f\|^2}$ and the proof is complete. □

5.4.1. *Duality for Hilbert space.* In the case $\mathbb{F} = \mathbb{R}$ the Riesz representation theorem identifies H^* with H. In the case $\mathbb{F} = \mathbb{C}$, the mapping sending $\lambda \in H^*$ to the vector h_0 is *conjugate* linear and thus H^* is not exactly H (under this map). However, it is customary when working in complex Hilbert space not to make this distinction. This convention creates some conflicts that must kept in mind. For instance, given Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ and a bounded linear map $T : \mathcal{X} \to \mathcal{Y}$, the *adjoint* of T, denoted T^* is the uniquely determined (by Hahn-Banach) linear map $T^* : \mathcal{Y}^* \to \mathcal{X}^*$ defined by $Tf = f \circ T$ so that $Tf(x) = f(T(x))$ for $x \in \mathcal{X}$. See Theorem 2.32 Because of our conjugate linear identification of H with H^*, the notion of the adjoint of a operator in the context of Hilbert space differs from the of operators between Banach spaces as described in the following proposition.

Proposition 5.29. *If H, K are Hilbert spaces and $T : H \to K$ is a bounded operator, then there is a unique bounded operator $S : K \to H$ satifying,*

$$\langle Th, k\rangle = \langle h, Sk\rangle.$$

Definition 5.30. The operator S associated to T in Proposition 5.29 is the *(Hilbert space) adjoint* of T, denoted T^* (sic).

The following elementary lemma will be used in the proof of Proposition 5.29 and elsewhere without comment.

Lemma 5.31. *Suppose H, K are Hilby spaces and $Y : H \to K$. If $\langle Th, k\rangle = 0$ for all $h \in H$ and $k \in K$, then $Y = 0$*

Proof. Given h choose $k = Th$ and use positive definitiness of the inner product. □

Proof. Define $S : K \to H$ as follows. Given $k \in K$, observe that the mapping $\lambda : H \to \mathbb{C}$ defined by $\lambda(f) = \langle Tf, k\rangle$ is (linear and) continuous. Hence, there is a vector Sk such that

$$\langle Tf, h\rangle = \lambda(f) = \langle f, Sk\rangle.$$

Conversely, if $S' : K \to H$ is linear and

$$\langle Tf, k\rangle = \langle f, S'k\rangle$$

for all $f, \in H$ and $k \in K$, then $\langle (S - S')k, f\rangle = 0$ for all $f \in H$ and $k \in K$ and hence $S' = S$. □

Further properties of the adjoints on Hilbert space appear in Problem 5.2.

A bounded operator T on a Hilbert space H is *self-adjoint* or *hermitian* if $T^* = T$.

Proposition 5.32. *If T is a bounded self-adjoint operator on a Hilbert space H, then $[\cdot, \cdot] : H \times H \to \mathbb{F}$ defined by $[f, g] = \langle Tf, g\rangle$ is a bilinear/sesquilinear form on H. If, in addition, $\langle Th, h\rangle = 0$ for all $h \in H$, then $T = 0$.*

Proof. Define $[\cdot, \cdot] : H \times H \to \mathbb{F}$ by

$$[f, g] = \langle Tf, g\rangle.$$

Since T is self-adjoint, $[g, f] = \langle Tg, f\rangle = \langle g, Tf\rangle = \overline{\langle Tf, g\rangle} = \overline{[f, g]}$, from which it follows that $[\cdot, \cdot]$ is a bilinear/sesquilinear form on H. Hence, by the polarization identity (Theorem 5.13),

$$4\langle Tf, g\rangle = \sum_{k=0}^{3} i^k \langle T(f + i^k g), f + i^k g\rangle = 0,$$

for all $f, g \in H$. By Lemma 5.31, $T = 0$. □

5.4.2. *Projections.* Returning to Theorem 5.19, if $M \leq H$ and $h \in H$, there exists a unique $f_0 \in M$ such that $(h - f_0) \perp M$. We thus obtain a well-defined function $P : H \to H$ (or, we could write $P : H \to M$) defined by

$$Ph = f_0. \tag{33}$$

That is, Ph is characterized by $Ph \in M$ and $(h - Ph) \in M\perp$. If the space M needs to be emphasized we will write P_M for P.

Definition 5.33. *A bounded operator Q on a Hilbert space H (meaning $Q : H \to H$ is linear and bounded) is a* projection *if $Q^* = Q$ and $Q^2 = Q$.* □

The following Theorem says if Q is a projection, then $Q = P_N$, where N is the range of Q; that is, Q is uniquely determined by its range, justifying the use of *the* in Definition 5.33; and conversely, if $M \le H$, then P_M is a projection (onto M).

Theorem 5.34. *Suppose $M \le H$. The mapping $P = P_M$ is a projection with range M. Moreover, if Q is a projection with range N, then*

(i) if $h \in N$, then $Qh = h$;
(ii) $\|Qh\| \le \|h\|$ for all $h \in H$;
(iii) $N \le H$;
(iv) $N^\perp$ is the kernel of Q;
(v) $I - Q$ is a projection with range $N^\perp$; and
(vi) $Q = P_N$.

Definition 5.35. For $M \le H$ and Q the operator P_M is called the *orthogonal projection of H on M* and, for $h \in H$, the vector $P_M h$ is the *orthogonal projection* of h onto M.

Proof. In view of Corollary 5.25, $M \oplus M^\perp = H$, from which it follows readily that P is a linear map.

Evidently P maps into M and if $f \in M$, then $Pf = f$ and hence P maps onto M and $PPf = Pf$ (and so $P^2 = P$).

If $h \in H$, then $h = Ph + (h - Ph)$. But $(h - Ph) \in M^\perp$ and $Ph \in M$, and thus, by the Pythagorean Theorem

$$\|h\|^2 = \|h - Ph\|^2 + \|Ph\|^2.$$

Hence $\|Ph\| \le \|h\|$. In particular, P is a bounded operator on H. (See also Problem 3.10.)

Given $g, f \in H$, since $g - Pg$ is orthogonal to M and Pf is in M,

$$\begin{aligned}\langle Pf, Pg\rangle =&\langle Pf, Pg\langle + \langle Pf, (g - Pg)\rangle \\ =&\langle Pf, g\rangle = \langle f, P^* g\rangle.\end{aligned}$$

On the other hand, by the same reasoning

$$\begin{aligned}\langle Pf, Pg\rangle =&\langle Pf + (I - P)f, Pg\rangle \\ =&\langle f, Pg\rangle.\end{aligned}$$

Hence $P^* = P$ and all the claims about P have now been proved.

Turning to Q, suppose Q is a projection and let N denote the range of Q. Since $Q^2 = Q$ it follows that $Qh = h$ for $h \in N$ (the range of Q). Also from $Q^2 = Q$ we have $Q(I - Q) = 0$. Thus if $h, f \in H$, then

$$\langle Qh, (I - Q)f\rangle = \langle h, Q(I - Q)f\rangle = 0.$$

Choosing $f = h$, it follows that $h = Qh + (I - Q)h$ is an orthogonal decomposition and hence $\|Qh\| \le \|h\|$ and so Q is continuous.

If (h_n) is a sequence from the range of Q that converges to $h \in H$, then, by continuity of Q, the sequence $(h_n = Qh_n)$ converges to Qh and thus $h = Qh$ so that the range of Q is closed.

Next, $f \in N^\perp$ if and only if

$$0 = \langle Qh, f\rangle = \langle h, Qf\rangle$$

for every $h \in H$; if and only if $Qf = 0$. Thus $N^\perp = \ker(Q)$.

An easy argument shows $I - Q$ is a projection too. In particular, f is in the range of $I - Q$ if and only if $(I - Q)f = f$. On the other hand $(I - Q)f = f$ if and only if $Qf = 0$. Thus the range of $I - Q$ is the kernel of Q. Finally, given $h \in H$, we have $Qh \in N$ and $h - Qh = (I - Q)h \in N^\perp$. Thus $Q = P_N$. □

END Friday 2025-02-21 Though we had not proved the assertions about Q in Proposition 5.29 nor Proposition 5.32.

5.5. **Orthonormal Sets and Bases.** Recall, a subset E of a Hilbert space H is *orthonormal* if $\|e\| = 1$ for all $e \in E$, and if $e, f \in E$ and $e \neq f$, then $e \perp f$.

Definition 5.36. An orthonormal set is *maximal* if it is not contained in any larger orthonormal set. A maximal orthonormal set is called an *(orthonormal) basis* or a *Hilbert space basis* for H.

Proposition 5.37. *An orthonormal set E is maximal if and only if $E^\perp = \{0\}$ if and only if the $\overline{\operatorname{span} E} = H$.*

An subset E of a Hilbert space H is a *complete orthonormal set* if E is orthonormal and $E^\perp = \{0\}$. Thus, E is a Hilbert space basis if and only if E is a complete orthonormal set.

Proof. Suppose E is not maximal. Hence there is an orthonormal set $F \supsetneq E$ and a vector $f \in F \setminus E$. In particular, $0 \neq f \in E^\perp$. Conversely, if $0 \neq f \in E^\perp$, then $F = E \cup \{\frac{f}{\|f\|}\}$ is an orthormal set that properly contains E and hence E is not maximal.

For the second part, from what has been proved, E is maximal if and only if $E^\perp = \{0\}$ if and only if $E^{\perp\perp} = H$. On the other hand $\overline{\operatorname{span} E} = E^{\perp\perp}$ by Corollary 5.25 □

Remark 5.38. It must be stressed that a basis in the above sense need *not* be a basis in the sense of linear algebra; that is, a basis for H as a vector space. In particular, it is always true that an orthonormal set is linearly independent (Exercise: prove this statement), but in general an orthonormal basis need not span H. In fact, if E is an infinite orthonormal subset of H, then E does not span H. See Problem 3.6.

If E is an orthonormal set in a Hilbert space H, then E is a basis for the Hilbert space span E.

Example 5.39. Here are some common examples of orthonormal bases.

(a) Of course the standard basis $\{e_1, \dots, e_n\}$ is an orthonormal basis of $\mathbb{F}^n$.

(b) In much the same way we get a orthonormal basis of $\ell^2(\mathbb{N})$; for each n define

$$e_n(k) = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n \end{cases}$$

It is straightforward to check that the set $E = \{e_n\}_{n=1}^{\infty}$ is orthonormal. In fact, it is a basis. To see this, notice that if $h : \mathbb{N} \to \mathbb{F}$ belongs to $\ell^2(\mathbb{N})$, then $\langle h, e_n \rangle = h(n)$, and hence if $h \perp E$, we have $h(n) = 0$ for all n, so $h = 0$.

(c) Let $H = L^2[0,1]$. Consider for $n \in \mathbb{Z}$ the set of functions $E = \{e_n(x) = e^{2i\pi nx} : n \in \mathbb{Z}\}$. An easy exercise shows this set is orthonormal. Though not obvious, it is in fact a basis. (See Problem 5.6.) Here is an outline of a proof. Given a (Lebesgue) measurable set $E \subseteq [0,1]$, by regularity there exists an open set U and a closed set F such that $F \subseteq E \subseteq U$ and $m(U \setminus F) < \epsilon$, where m is Lebesgue measure. Let $K = [0,1] \setminus U$ and define $f : [0,1] \to \mathbb{R}$ by

$$f(t) = \frac{d(t,K)}{d(t,F) + d(t,K)},$$

where $d(t,S) = \inf\{|t-s| : s \in S\}$ is the distance from a point t to the set $S \subseteq [0,1]$. Because F, K are compact, the infima in these distance are attained. In particular, $f(t) = 1$ for $t \in F$, and $f(t) = 0$ for $t \in K$, while otherwise $0 \le f(t) \le 1$. It follows that

$$\int_0^1 |f - \chi_E|^2 \, dm \le \mu(U \setminus F) < \epsilon,$$

Since simple functions are dense in $L^2([0,1])$ (an exercise), it follows that continuous functions are too. Stone Weierstrass implies that the span of E (the set of trigonmetric polynomials) is *uniformly* dense in $C([0,1])$.

END Monday 2025-02-24

Theorem 5.40. *Every Hilbert space $H \neq \{0\}$ has an orthonormal basis.*

Proof. The proof is essentially the same as the Zorn's lemma proof that every (nontrivial) vector space has a basis. Let H be a Hilbert space and $\mathcal{E}$ the collection of orthonormal subsets of H, partially ordered by inclusion. Since $H \neq (0)$, the collection $\mathcal{E}$ is not empty. If (E_α) is an ascending chain in $\mathcal{E}$, then it is straightforward to verify that $\cup_\alpha E_\alpha$ is an orthonormal set, and is an upper bound for (E_α). Thus by Zorn's lemma, $\mathcal{E}$ has a maximal element, say E. ∎

Proposition 5.41 (Bessel's Inequality)**.** *If E is an orthornormal set in a Hilbert space H, then, for each $h \in H$,*

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

In particular, $E_h = \{e \in E : \langle h, e\rangle \neq 0\}$ is at most countable.

Proof. For a finite subset F of E, observe that h is the sum of the orthogonal vectors $f = \sum_{e\in F}\langle h, e\rangle e$ and $h - f$. Hence,

$$\|h\|^2 = \|f\|^2 + \|h - f\|^2 \geq \|f\|^2 = \sum_{e\in F} |\langle h, e\rangle|^2.$$

Thus,

$$\|h\|^2 \geq \sup\{\sum_{e\in F} |\langle h, e\rangle|^2 : F \subseteq E, |F| < \infty\} = \sum_{e\in E} |\langle h, e\rangle|^2.$$ □

5.6. **Convergent series in Hilbert space and basis expansions.** This section begins with a discussion of convergence of infinite series in Hilbert space before turning to basis expansions and Parseval's equality.

We have already encountered ordinary convergence and absolute convergence in our discussion of completeness: recall that the series $\sum_{n=1}^{\infty} h_n$ *converges* if $\lim_{N\to\infty} \sum_{n=1}^{N} h_n$ exists; its limit h is called the sum of the series. The series *converges absolutely* if $\sum_{n=1}^{\infty} \|h_n\| < \infty$ and absolute convergence implies convergence.

Definition 5.42. Suppose H is a Hilbert space and $(h_n)_{n=1}^{\infty}$ is a sequence from H. The series $\sum_{n=1}^{\infty} h_n$ is *unconditionally convergent* if there exists an $h \in H$ such that for each bijection $\varphi : \mathbb{N} \to \mathbb{N}$ the series $\sum_{n=1}^{\infty} h_{\varphi(n)}$ converges to h. (In other words, every reordering of the series converges, and to the same sum.)

Remark 5.43. Of course absolute convergence implies unconditional convergence. For ordinary scalar series, or in a finite dimensional Hilbert space such as $\mathbb{F}^n$, unconditional convergence implies absolute convergence; however in infinite dimensional Hilbert space unconditional convergence need not imply absolute convergence as example 5.45 following the proof of Theorem 5.44 shows.

Theorem 5.44. *Suppose $E = \{e_1, e_2, \dots\} \subseteq H$ is a countable orthonormal set and (a_n) is a sequence of complex numbers. The following are equivalent.*

(i) the series $\sum_{j=1}^{\infty} a_j e_j$ converges;
(ii) $\sum_{j=1}^{\infty} |a_j|^2$ converges; and
(iii) the series $\sum_{j=1}^{\infty} a_j e_j$ converges unconditionally.

If $\sum_{j=1}^{\infty} a_j e_j$ converges to g, then $\langle g, e_j\rangle = a_j$ for all j.

Further, if $h \in H$, then the series

$$\sum_{j=1}^{\infty} \langle h, e_j \rangle e_j \tag{34}$$

is unconditionally convergent and, letting g denote the (unconditional) sum,

$$\langle g, e_j \rangle = \langle h, e_j \rangle$$

for all j.

Proof. Let s_n denote the partial sums of the series $\sum_{j=1}^{\infty} a_j\, e_j$,

$$s_n = \sum_{j=1}^{n} a_j\, e_j.$$

Since H is complete, the series $\sum_{j=1}^{\infty} a_j e_j$ converges (meaning (s_n) converges) if and only if for each $\epsilon > 0$ there is an N so that for all $m \geq n \geq N$,

$$\|s_m - s_n\|^2 = \sum_{j=n+1}^{m} |a_j|^2 < \epsilon \tag{35}$$

(meaning (s_n) is Cauchy) if and only if the series $\sum_{j=1}^{m} |a_j|^2$ converges. Hence items (i) and (ii) are equivalent.

Now suppose $s_n = \sum_{j=1}^{n} a_j e_j$ converges to, say, g and $\varphi : \mathbb{N} \to \mathbb{N}$ is a permutation (bijection). Let $s'_n = \sum_{j=1}^{n} a_{\varphi(j)} e_{\varphi(j)}$. Given $\epsilon > 0$, choose N so that (35) holds. In particular,

$$\sum_{j=N+1}^{\infty} |a_j|^2 \leq \epsilon.$$

Now choose $M \geq N$[3] so that

$$J_N = \{1, 2, \dots N\} \subseteq \{\varphi(1), \varphi(2), \dots \varphi(M)\}.$$

For $n \geq M$ let $J_n = \{1, \dots n\}$ and $J'_n = \{\varphi(1), \dots \varphi(n)\}$ and let G_n denote their symmetric difference; that is $G_n = (A_n \setminus B_n) \cup (B_n \setminus A_n)$. From

$$s_n - s'_n = \sum_{j \in J_n} a_j e_j \sum_{j \in J'_n} = \sum_{j \in J_n \setminus J'_n} a_j e_j - \sum_{j \in J'_n \setminus J_n} a_j e_j,$$

it follows that

$$\|s_n - s'_n\|^2 = \sum_{k \in G_n} \|a_k\|^2.$$

On the other hand $G_n \subseteq J_N^c$, since $J_N \subseteq J_n, J'_n$. Therefore,

$$\|s_n - s'_n\|^2 = \sum_{k \in G_n} |a_k|^2 \leq \sum_{N+1}^{\infty} |a_k|^2 \leq \epsilon.$$

[3]For instance $M = \max \varphi^{-1}(J_N)$.

Hence (s'_n) converges to g too. Hence item (ii) implies item (iii) and the proof of the first part of the theorem is complete, since evidently item (iii) implies item (i).

Now suppose $\sum_{j=1}^{\infty} a_j e_j$ converges to g and set $s_n = \sum_{j=1}^{n} a_j e_j$. Using Lemma 5.4, since (s_n) converges to g, for each m, the sequence $(\langle s_n, e_m\rangle)_n$ converges to $\langle g, e_m\rangle$. On the other hand, $\langle s_n, e_m\rangle = a_m$ for $n \geq m$. Hence $\langle g, e_m\rangle = a_m$.

For $h \in H$ Bessel's inequality, Theorem 5.41, implies the convergence of $\sum |\langle h, e_j\rangle|^2$ and thus, by what has already been proved, the series $\sum \langle h, e_j\rangle\, e_j$ converges (unconditionally) to some $g \in H$ and $\langle g, e_m\rangle = \langle h, e_m\rangle$ for all m. □

Example 5.45. Suppose $\{e_1, e_2, \dots\}$ is a countable orthonormal set in a Hilbert space H. The series

$$\sum_{j=1}^{\infty} \frac{1}{j} e_j$$

is Cauchy (verify this as an exercise) and hence converges to some $h \in H$. From Theorem 5.44 it follows that $\langle h, e_j\rangle = \frac{1}{j}$ and the series above converges unconditionally to h. On the other hand, this series does not converge absolutely and hence unconditional convergence does not imply absolute convergence.

END Wednesday 2025-02-26

There is another notion of convergence in Hilbert space.

Definition 5.46. Suppose $\varnothing \neq S \subseteq H$ and let $\mathcal{F}$ denote the collection of finite subsets of S. The series

$$\sum_{s\in S} s$$

converges as a net if there exists $h \in H$ such that for every $\epsilon > 0$ there exists an $F \in \mathcal{F}$ such that for every $F \subseteq G \in \mathcal{F}$,

$$\|\sum_{s\in G} s - h\| < \epsilon.$$

Often S is presented as an indexed set, so that $S = \{h_i : i \in I\}$ for some set I, in which case the series is written as $\sum_{i\in I} h_i$.

Proposition 5.47. *If E is an orthonormal subset of a Hilbert space H and $h \in H$, then the series*

$$\sum_{e\in E} \langle h, e\rangle e$$

converges (as a net). Moreover, if g is the limit (as a net) of this series, then, for each $e \in E$,

$$\langle g, e\rangle = \langle h, e\rangle.$$

Proof. Let $E_h = \{e \in E : \langle h, e\rangle \neq 0\}$. From Bessel's inequality, Proposition 5.41, E_h is at most countable. Suppose E_h is countable and choose an enumeration, $E_h = \{e_1, e_2, \dots\}$. By Theorem 5.44, the series

$$\sum_{j=1}^{\infty} \langle h, e_j\rangle e_j$$

converges unconditionally to some $g \in H$ and moreover $\langle g, e_j\rangle = \langle h, e_j\rangle$ for all j. On the other hand, since the partial sums $s_n = \sum_{j=1}^{n} \langle h, e_j\rangle e_j$ converge to h (in norm), for each $e \in E \setminus E_h$, the sequence $(0 = \langle s_n, e\rangle)_n$ converges to $\langle g, e\rangle$ and so $\langle g, e\rangle = 0$. Hence $\langle g, e\rangle = \langle h, e\rangle$ for all $e \in E$. In particular, $E_g = E_h$.

To prove the series $\sum_{e\in E} \langle h, e\rangle e$ converges to g as a net, let $\epsilon > 0$ be given. There is an N so that

$$\|g - \sum_{j=1}^{N} \langle h, e_j\rangle e_j\| < \epsilon$$

and hence

$$\sum_{j=N+1}^{\infty} |\langle h, e_j\rangle|^2 < \epsilon^2.$$

Let $F = \{e_1, \dots, e_N\}$. If $G \subseteq E$ is finite and $F \subseteq G$, then, letting $T = G \setminus F$,

$$\begin{aligned} \|\sum_{e\in T} \langle h, e\rangle e\|^2 &= \|\sum_{e\in E_h \cap T} \langle h, e\rangle e + \sum_{e\in (E\setminus E_h)\cap T} \langle h, e\rangle e\|^2 \\ &= \|\sum_{e\in E_h\setminus F} \langle h, e\rangle e\|^2 \leq \sum_{j=N+1}^{\infty} |\langle h, e_j\rangle|^2 < \epsilon^2. \end{aligned}$$

Hence

$$\begin{aligned} \|g - \sum_{e\in G} \langle h, e\rangle\, e\| &\leq \|g - \sum_{e\in F} \langle h, e\rangle\| + \|\sum_{e\in T} \langle h, e\rangle\| \\ &= \|g - \sum_{j=1}^{N} \langle g, e_j\rangle\, e_j\| + \|\sum_{e\in T} \langle g, e\rangle e\| < 2\epsilon. \end{aligned}$$

Hence $\sum_{e\in E} \langle h, e\rangle$ converges as a net to g. $\square$

Item (d) in the following theorem is known as *Parseval's equality*. For each $h \in H$, the series in item (b) converges (as a net) to some $g \in H$ by Proposition 5.47.

Theorem 5.48. *If $E \subseteq H$ is an orthonormal set, then the following are equivalent:*

(a) E is a (orthonormal) basis for H;

(b) $h = \sum_{e\in E} \langle h, e\rangle e$ for each $h \in H$;

(c) $\langle g, h\rangle = \sum_{e\in E} \langle g, e\rangle\langle e, h\rangle$ for each $g, h \in H$; and

(d) $\|h\|^2 = \sum_{e\in E} |\langle h, e\rangle|^2$ for each $h \in H$.

Proof. Suppose E is an orthonormal set in H, but item (b) does not hold. Thus there is an $h \in H$ such that $h \neq \sum_{e\in E}\langle h, e\rangle e$. By Proposition 5.47,

$$\sum_{e\in E}\langle h, e\rangle e$$

converges (as a net) to some $g \in H$ and moreover $\langle g, e\rangle = \langle h, e\rangle$ for all $e \in E$. By assumption, $f = g - h \neq 0$. On the other hand,

$$\langle f, e\rangle = \langle g - h, e\rangle = 0,$$

and thus $E^{\perp} \neq \{0\}$ so that, by Proposition 5.37, E is not maximal. Hence item (a) implies item (b).

Now suppose item (b) holds and let $h, g \in H$ be given. Given ϵ, choose a finite subset F of E such that if $F \subseteq G \subseteq E$, then

$$\|h - \sum_{e\in G}\langle h, e\rangle e\|,\ \|g - \sum_{e\in G}\langle g, e\rangle e\| < \sqrt{\epsilon}$$

and observe, using the Cauchy-Schwarz inequality,

$$\begin{aligned}\epsilon >& \big|\langle h - \sum_{e\in G}\langle h, e\rangle e, g - \sum_{f\in G}\langle g, f\rangle f\big| \\ =& \big|\langle h, g\rangle - \sum_{e\in G}\langle h, e\rangle \langle e, g\rangle\big|.\end{aligned}$$

Hence item (b) implies item (c).

Item (d) follows from item (c) by choosing $g = h$. Finally, suppose that item (a) does not hold. In that case there exists a unit vector $h \in H$ such that h is orthogonal to E. Thus $\langle h, e\rangle = 0$ for all $e \in E$ so that

$$\sum_{e\in E} |\langle h, e\rangle|^2 = 0 \neq 1 = \|h\|^2$$

and item (d) does not hold. □

Given a set E, let V denote the vector space of finite linear combinations of elements of E and define an inner product on V by declaring $\langle e, f\rangle = 0$ if $e, f \in E$ and $e \neq f$ and $\langle e, e\rangle = 1$ for $e \in E$. The completion H of V (see Proposition 2.31) is a Hilbert space, denoted $\ell^2(E)$.

Corollary 5.49. *If E is a basis for a Hilbert space H, then H is isomorphic, as a Hilbert space, to $\ell^2(E)$.*

5.7. **Gram-Schmidt and Hilbert space dimension.**

Theorem 5.50. *Let $\{e_1, \dots e_n\}$ be an orthonormal set in H, and let $M = \operatorname{span}\{e_1, \dots e_n\}$. The orthogonal projection $P = P_M$ onto M is given by, for $h \in H$,*

$$Ph = \sum_{j=1}^{n} \langle h, e_j \rangle e_j. \tag{36}$$

Proof. Given $h \in H$, let $g = \sum_{j=1}^{n} \langle h, e_j \rangle e_j$. Since $g \in M$, it suffices to show $(h-g) \perp M$. For $1 \le m \le n$,

$$\begin{aligned} \langle h-g, e_m \rangle =& \langle h, e_m \rangle - \left\langle \sum_{j=1}^{n} \langle h, e_j \rangle e_j, e_m \right\rangle \\ &= \langle h, e_m \rangle - \sum_{j=1}^{n} \langle h, e_j \rangle \langle e_j, e_m \rangle \\ &= \langle h, e_m \rangle - \langle h, e_m \rangle = 0. \end{aligned}$$

It follows that $h - g$ is orthogonal to $\{e_1, \dots, e_n\}$ and hence to M. □

END Friday 2025-02-28

Theorem 5.51 (Gram-Schmidt process)**.** *If $(f_n)_{n=1}^{\infty}$ is a linearly independent sequence in H, then there exists an orthonormal sequence $(e_n)_{n=1}^{\infty}$ such that $\operatorname{span}\{f_1, \dots f_n\} = \operatorname{span}\{e_1, \dots e_n\}$ for each n.*

Proof. The proof proceeds by induction. Put $e_1 = f_1/\|f_1\|$ and note $\{e_1\}$ is an orthormal set and $\operatorname{span}\{e_1\} = \operatorname{span}\{f_1\}$. Assuming $e_1, \dots e_n$ have been constructed satisfying the conditions of the theorem, let $P = P_{M_n}$ where $M_n = \operatorname{span}\{e_1, \dots, e_n\}$ and let $g_{n+1} = f_{n+1} - \sum_{j=1}^{n} \langle f_{n+1}, e_j \rangle e_j = f_{n+1} - Pf$. By Theorem 5.50 g_{n+1} is orthogonal to M_n. It is also not 0 by the independence assumption on the f_j. Let $e_{n+1} = \frac{g_{n+1}}{\|g_{n+1}\|}$. □

Corollary 5.52. *Suppose H is a Hilbert space. If H has finite dimension $n \ge 1$ as a vector space, then there exists an orthonormal set $\{e_1, \dots, e_n\}$ in H that spans H. Conversely, if there is a positive integer n and an orthonormal set $\{e_1, \dots, e_n\}$ that spans H, then H has finite dimension n as a vector space.*

In particular, if H contains a finite maximal orthonormal set, then every maximal orthornormal set in E has the same cardinality and moreover this cardinality is the dimension of H as a vector space.

Remark 5.53. If H has a finite orthonormal basis $E = \{e_1, \dots, e_n\}$, then by Theorem 5.48(b), E spans (in the sense of linear algebra) and is therefore a vector space (Hamel) basis for H. Hence H has dimension n as a vector space and further every orthonormal basis of H has exactly n elements.

On the other hand, if H has an infinite orthonormal basis E, then it contains an infinite linearly independent set (the basis E) and so has infinite dimension as a vector space. □

Theorem 5.54. *Any two bases of a Hilbert space H have the same cardinality.*

The proof uses some basic facts about cardinality. Two sets A and B have the same cardinality, written $|A| = |B|$ if there is a bijection $f : A \to B$. If there is a one-one map $f : A \to B$ we write $|A| \leq |B|$. By the Cantor–Schröder–Bernstein theorem, if $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$. If A is an infinite set, then $|A \times \mathbb{N}| = |A|$[4] and if A is an infinite set and B_a is an at most countable set for each $a \in A$, then $|\cup_{a\in A} B_a| \leq |A \times B| = |A|$. The theorem says if E, F are orthonormal bases for H, then $|E| = |F|$.

Proof. Suppose E, F are orthonormal bases for H. If E is finite, then E is a basis in the vector space sense and thus H is finite dimensional as a vector space. Since F is orthonormal, it is linearly independent and hence $|F| \leq |E|$. Thus F is also a vector space basis for H and so $|F| = |E|$. By symmetry, either both E and F are finite and have the same cardinality or both are infinite. Accordingly suppose both are infinite.

Fix $e \in E$ and consider the set

$$F_e = \{f \in F \mid \langle f, e\rangle \neq 0\}.$$

Since F is orthonormal, each F_e is at most countable by Proposition 5.41 and since E is a basis, each $f \in F$ belongs to at least one F_e. Thus $\bigcup_{e\in E} F_e = F$, and

$$|F| = \left|\bigcup_{e\in E} F_e\right| \leq |E \times \mathbb{N}| = |E|,$$

where the last equality holds since E is infinite.

By symmetry, $|E| \leq |F|$ and the proof is complete. □

In light of this theorem, we make the following definition.

Definition 5.55. The *(orthogonal) dimension* of a Hilbert space H is the cardinality of any orthonormal basis, and is denoted $\dim H$. If $\dim H$ is finite or countable, H is *separable* and in this case the terminology H is a *separable Hilbert space* is commonly used.

Corollary 5.56. *Suppose H is a Hilbert space. If H is finite dimensional as a vector space, then H is separable as a metric space.*

---

[4] That $|S \times S| = |S|$ for an infinite set S in full generality requires the axiom of choice. On the other hand, since $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}$ the proof of Theorem 5.54 given below shows if H has a countable orthonormal basis, then all orthonormal basis of H are countable.

If H is not finite dimensional as a vector space, then H is separable as a metric space if and only if there is a a countable orthonormal set $E = \{e_1, e_2, \dots\}$ such that $\overline{\operatorname{span} E} = H$.

Proof. We consider the complex case of $\mathbb{F} = \mathbb{C}$, the real case being similar. If H has a countable orthonormal basis $E = \{e_1, e_2, \dots\}$, then the set $D = \{\sum_{j=1}^n a_j e_j : a_j \in \mathbb{Q} + i\mathbb{Q},\, n \in \mathbb{N}\}$ is dense in H since $E \subseteq D$ so that $H = \overline{\operatorname{span} E} \subseteq \overline{\operatorname{span} D} \subseteq H$.

The proof that a finite dimensional Hilbert space is separable as a metric space is similar to the proof above.

Now suppose a basis of E contains uncountably many elements (and thus all bases of E are uncountable by Theorem 5.54). Since $\|e - f\| = \sqrt{2}$ for all $e, f \in H$ such that $e \neq f$, if C is a countable subset of H, then $E \not\subseteq \cup_{c \in C} B(c, 1)$ and hence C is not dense. Thus H is not separable as a metric space. □

5.8. **Weak convergence.** [Optional] In addition to the norm topology, Hilbert spaces carry another topology called the *weak topology*. In these notes we will stick to the seperable case and just study weakly convergent sequences.

Definition 5.57. Let H be a seperable Hilbert space. A sequence (h_n) in H *converges weakly to* $h \in H$ if for all $g \in H$,

$$\langle h_n, g \rangle \to \langle h, g \rangle.$$

The Cauchy-Schwarz inequality implies if (h_n) converges to h in norm, then (h_n) converges weakly to h. However, when H is infinite-dimensional, the converse can fail. For instance, let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal basis for H. Then (e_n) converges to 0 weakly. (The proof is an exercise, see Problem 5.9). On the other hand, the sequence (e_n) is not norm convergent, since it is not Cauchy. In this section weak convergence is characterized as "bounded coordinate-wise convergence" and it is shown that the unit ball of a separable Hilbert space is weakly sequentially compact.

Proposition 5.58. *Let H be a Hilbert space with orthonormal basis $\{e_j\}_{j=1}^{\infty}$. A sequence (h_n) in H is weakly convergent if and only if*

i) $\sup_n \|h_n\| < \infty$, *and*
ii) $\lim_n \langle h_n, e_j \rangle$ *exists for each j.*

Proof. Suppose (h_n) converges to h weakly. For each n

$$L_n(g) = \langle g, h_n \rangle$$

is a bounded linear functional on H. Since, for fixed g, the sequence $|L_n(g)|$ converges, it is bounded. Thus, the family of linear functionals (L_n) is pointwise bounded and hence, by the Principle of Uniform boundedness, $\sup \|h_n\| = \sup \|L_n\| < \infty$, showing (i) holds. Item (ii) is immediate from the definition of weak convergence.

Conversely, suppose (i) and (ii) hold, let $M = \sup \|h_n\|$. Define

$$\hat{h}_j = \lim \langle h_n, e_j \rangle.$$

We will show that $\sum_j |\hat{h}_j|^2 \le M$ (so that the series $\sum \hat{h}_j e_j$ is norm convergent in H); we then define h to be the sum of this series and show that $h_n \to h$ weakly.

For positive integers J and all n,

$$\sum_{j=1}^{J} |\langle h_n, e_j \rangle|^2 \le \|h_n\|^2 \le M^2$$

by Bessel's inequality. Thus,

$$\sum_{j=1}^{J} |\hat{h}_j|^2 = \sum_{j=1}^{J} \lim_n |\langle h_n, e_j \rangle|^2 = \lim_n \sum_{j=1}^{J} |\langle h_n, e_j \rangle|^2 \le M^2.$$

Thus $\sum_j |\hat{h}_j|^2 \le M^2$ and therefore the series $\sum_j \hat{h}_j e_j$ is norm convergent to some $h \in H$ such that $\langle h, e_j \rangle = \hat{h}_j$ by Theorem 5.44. By Theorem 5.48, $\|h\| \le M$.

Now we prove that (h_n) converges to h weakly. Fix $g \in H$ and let $\epsilon > 0$ be given. Since $g = \sum_j \langle g, e_j \rangle e_j$ (where the series is norm convergent) there exists an positive integer J large enough so that

$$\left\| g - \sum_{j=1}^{J} \langle g, e_j \rangle e_j \right\| = \left\| \sum_{j=J+1}^{\infty} \langle g, e_j \rangle e_j \right\| < \epsilon.$$

Let $g_0 = \sum_{j=1}^{J} \langle g, e_j \rangle e_j$, write $g = g_0 + g_1$, observe $\|g_1\| < \epsilon$ and estimate,

$$|\langle h_n - h, g \rangle| \le |\langle h_n - h, g_0 \rangle| + |\langle h_n - h, g_1 \rangle|.$$

By (ii), the first term on the right hand side goes to 0 with n, since g_0 is a finite sum of e_j's. By Cauchy-Schwarz, the second term is bounded by $2M\epsilon$. As ϵ was arbitrary, we see that the left-hand side goes to 0 with n. □

It turns out, if (h_n) converges to h weakly, then $\|h\| \le \liminf \|h_n\|$ and further, still assuming (h_n) converges weakly to h, $\|h\| = \lim \|h_n\|$ if and only if (h_n) converges to h in norm. See Problem 5.9.

Theorem 5.59 (Weak compactness of the unit ball in Hilbert space)**.** *If (h_n) is a bounded sequence in a separable Hilbert space H, then (h_n) has a weakly convergent subsequence.*

Proof. Using the previous proposition, it suffices to fix an orthonromal basis (e_j) and produce a subsequence $(h_{n_k})_k$ such that $\langle h_{n_k}, e_j \rangle$ converges for each j. This is a standard "diagonalization" argument, and the details are left as an exercise (Problem 5.11) □

5.9. **Problems.**

Problem 5.1. Prove the complex form of the polarization identity: if H is a Hilbert space over $\mathbb{C}$, then for all $g, h \in H$

$$\langle g, h\rangle = \frac{1}{4}\left(\|g+h\|^2 - \|g-h\|^2 + i\|g+ih\|^2 - i\|g-ih\|^2\right)$$

Problem 5.2. (Adjoint operators) Let H be a Hilbert space and $T : H \to H$ a bounded linear operator.

a) Prove there is a unique bounded operator $T^* : H \to H$ satisfying $\langle Tg, h\rangle = \langle g, T^*h\rangle$ for all $g, h \in H$, and $\|T^*\| = \|T\|$.
b) Prove, if $S, T \in B(H)$, then $(aS+T)^* = \overline{a}S^* + T^*$ for all $a \in \mathbb{F}$, and that $T^{**} = T$.
c) Prove $\|T^*T\| = \|T\|^2$.
d) Prove $\ker T$ is a closed subspace of H, $\overline{(\mathrm{ran} T)} = (\ker T^*)^\perp$ and $\ker T^* = (\mathrm{ran} T)^\perp$.

Problem 5.3. Let H, K be Hilbert spaces. A linear transformation $T : H \to K$ is called *isometric* if $\|Th\| = \|h\|$ for all $h \in H$, and *unitary* if it is a surjective isometry. Prove the following:

a) T is an isometry if and only if $\langle Tg, Th\rangle = \langle g, h\rangle$ for all $g, h \in H$, if and only if $T^*T = I$ (here I denotes the identity operator on H).
b) T is unitary if and only if T is invertible and $T^{-1} = T^*$, if and only if $T^*T = TT^* = I$.
c) Prove, if $E \subseteq H$ is an orthonormal set and T is an isometry, then $T(E)$ is an orthonormal set in K.
d) Prove, if H is finite-dimensional, then every isometry $T : H \to H$ is unitary.
e) Consider the *shift operator* $S \in B(\ell^2(\mathbb{N}))$ defined by

$$S(a_0, a_1, a_2, \dots) = (0, a_0, a_1, \dots) \tag{37}$$

Prove S is an isometry, but not unitary. Compute S^* and SS^*.

Problem 5.4. For any set J, let $\ell^2(J)$ denote the set of all functions $f : J \to \mathbb{F}$ such that $\sum_{j\in J} |f(j)|^2 < \infty$. Then $\ell^2(J)$ is a Hilbert space.

a) Prove $\ell^2(I)$ is isometrically isomorphic to $\ell^2(J)$ if and only if I and J have the same cardinality. (Hint: use Problem 5.3(c).)
b) Prove, if H is any Hilbert space, then H is isometrically isomorphic to $\ell^2(J)$ for some set J.

Problem 5.5. Let $(X, \mathscr{M}, \mu)$ be a σ-finite measure space. Prove the simple functions that belong to $L^2(\mu)$ are dense in $L^2(\mu)$.

Problem 5.6. (The Fourier basis) Prove the set $E = \{e_n(t) := e^{2\pi int} | n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2[0, 1]$. (Hint: use the Stone-Weierstrass theorem to prove that the set of trigonometric polynomials $P = \{\sum_{n=-M}^{N} c_n e^{2\pi int}\}$ is uniformly dense in the space of continuous functions f on $[0,1]$ that satisfy $f(0) = f(1)$. Then show that this space of continuous functions is dense in $L^2[0,1]$. Finally show that if f_n is a sequence in $L^2[0,1]$ and $f_n \to f$ uniformly, then also $f_n \to f$ in the L^2 norm.)

Problem 5.7. Let $(g_n)_{n\in\mathbb{N}}$ be an orthonormal basis for $L^2[0,1]$, and extend each function to $\mathbb{R}$ by declaring it to be 0 off of $[0,1]$. Prove the functions $h_{mn}(x) := 1_{[m,m+1]}(x)g_n(x-m)$, $n \in \mathbb{N}, m \in \mathbb{Z}$ form an orthonormal basis for $L^2(\mathbb{R})$. (Thus $L^2(\mathbb{R})$ is separable.)

Problem 5.8. Let $(X, \mathscr{M}, \mu), (Y, \mathscr{N}, \nu)$ are σ-finite measure spaces, and let $\mu\times\nu$ denote the product measure. Prove, if (f_m) and (g_n) are orthonormal bases for $L^2(\mu), L^2(\nu)$ respectively, then the collection of functions $\{h_{mn}(x,y) = f_m(x)g_n(y)\}$ is an orthonromal basis for $L^2(\mu \times \nu)$. Use this result to construct an orthonormal basis for $L^2(\mathbb{R}^n)$, and conclude that $L^2(\mathbb{R}^n)$ is separable.

Problem 5.9. (Weak Convergence)

a) Prove, if (h_n) converges to h in norm, then also (h_n) converges to h weakly. (Hint: Cauchy-Schwarz.)
b) Prove, if H is infinite-dimensional, and (e_n) is an orthonormal sequence in H, then $e_n \to 0$ weakly, but $e_n \not\to 0$ in norm. (Thus weak convergence does not imply norm convergence.)
c) Prove (h_n) converges to h in norm if and only if (h_n) converges to h weakly and $\|h_n\| \to \|h\|$.
d) Prove if (h_n) converges to h weakly, then $\|h\| \le \liminf \|h_n\|$.

Problem 5.10. Suppose H is countably infinite-dimensional (separable Hilbert space). Prove, if $h \in H$ and $\|h\| < 1$, then there is a sequence h_n in H with $\|h_n\| = 1$ for all n, and (h_n) converges to h weakly, but h_n does not converge to h strongly.

Problem 5.11. Prove Theorem 5.59.

Problem 5.12. Prove, if (a_n) is a sequence of complex numbers, then the following are equivalent.

(1) $\sum_{n\in\mathbb{N}} a_n$ converges as a net;
(2) $\sum_{n=1}^\infty a_n$ converges unconditionally;
(3) $\sum_{n=1}^\infty a_n$ converges absolutely.

Problem 5.13. Suppose (h_n) is a sequence from a Hilbert space H. Show, if $\sum_{n=1}^\infty h_n$ converges absolutely, then $\sum_{n=1}^\infty h_n$ converges unconditionally and as a net.

Problem 5.14. Suppose H is a Hilbert space and (h_j) is a sequence from H. Show, $\sum_{j=1}^\infty h_j$ converges unconditionally if and only if $\sum_{j\in\mathbb{N}} h_j$ converges as a net. (Warning: showing unconditional convergence implies convergence as a net is challenging.)

6. Signed measures

In this section we consider measures with codomain $\mathbb{F}$ (either $\mathbb{R}$ of $\mathbb{C}$) instead of $[0,\infty]$.

6.1. **Definitions, examples and elementary properties.**

Definition 6.1. Let $(X,\mathscr{M})$ be a measurable space. A *signed measure* or an $\mathbb{F}$-*measure* is a countably additive function $\rho:\mathscr{M}\to\mathbb{F}$; that is, if $(E_n)_{n=1}^\infty$ is a disjoint sequence of measurable sets, then

$$\sum_{n=1}^{\infty}\rho(E_n)=\rho(\bigcup_{n=1}^{\infty}E_n). \tag{38}$$

Sometimes the terminology positive measure is used instead of simply measure to indicate ρ takes values in $[0,\infty]$ and then *finite positive measure* indicates ρ takes values in $[0,\infty)$.

Remark 6.2. Several remarks are in order before proceeding.

(a) Choosing $E_n=\varnothing$ for all n obtains $\rho(\varnothing)=\sum_{n=1}^\infty\rho(\varnothing)$. Hence $\rho(\varnothing)=0$.

(b) Since $\rho(\varnothing)=0$, it follows that the countable additivity condition also includes finite additivity by choosing $E_n=\varnothing$ as needed to pass from a finite set of sets to a countable one.

(c) Given a permutation $\pi:\mathbb{N}\to\mathbb{N}$, since $\rho(\bigcup_{n=1}^\infty E_n)=\rho(\bigcup_{n=1}^\infty E_{\pi(n)})$, the series in equation (38) converges unconditionally and hence, by Riemann's rearrangement theorem, absolutely.

Alternately, one can see directly that the series in equation (38) must converges absolutely as follows. First take $\mathbb{F}=\mathbb{R}$ and let Let $K_+=\{k:\rho(E_k)\ge 0\}$ and let $K_-=\{k:\rho(E_k)<0\}$. The collections $A_\pm=\{E_k:k\in K_\pm\}$ are at most countable and their elements are pairwisde disjoint. Hence,

$$\pm\rho(\cup_{k\in K_\pm}E_k)=\pm\sum_{k\in K_\pm}\rho(E_k)=\sum_{k\in K_\pm}|\rho(E_k)|,$$

where the fact that the order of summation is immaterial for series with non-negative terms has been used. The complex case follows from the real case, since real ρ and image ρ are both real-measures and $|\rho(E)|\le|\operatorname{real}\rho(E)|+|\operatorname{image}\rho(E)|$.

(d) The argument in item (c) also proves the following in the case $\mathbb{F}=\mathbb{R}$. Given disjoint sets $E_1,\dots,E_n\in\mathscr{M}$, let $A=\cup\{E_k:\rho(E_k)\ge 0\}$ and $B=\cup\{E_k:\rho(E_k)<0\}$ and observe

(i) $\rho(A),-\rho(B)\ge 0$;

(ii) $E=\cup E_j=A\cup B$;

(iii) $\rho(E) = \rho(A) - \rho(B)$; and

(iv) $\sum |\rho(E_j)| = |\rho(A)| + |\rho(B)| = \rho(A) - \rho(B)$.

(e) In the case $\mathbb{F} = \mathbb{R}$ the theory can be developed allowing ρ to take values in either $(-\infty, \infty]$ or $[-\infty, \infty)$ (so as to avoid $\infty - \infty$). We will eschew this extra generality.

(f) If $\rho_1, \dots, \rho_n$ are finite positive measures on a measure space $(X, \mathscr{M})$ and $a_1, \dots, a_n \in \mathbb{F}$, then $\rho = \sum_{j=1}^{n} a_j \rho_j$ is an $\mathbb{F}$-measure on $\mathscr{M}$. In this way $\mathbb{M}_{\mathbb{F}}(\mathscr{M})$, the set of measures on $(X, \mathscr{M})$ becomes a vector space.

(g) If μ is a finite positive measure on a measure space $(X, \mathscr{M})$ and $F \in \mathscr{M}$, then

$$\rho(E) = \mu(E \cap F) - \mu(E \cap F^c)$$

defines an $\mathbb{R}$-measure on $\mathscr{M}$.

END Monday 2025-03-03, though we had not yet discussed items (d), (e) and (f) in the remark above

Proposition 6.3. *If $(X, \mathscr{M}, \mu)$ is a measure space and $f \in \mathcal{L}^1(\mu)$, then the function $\mu_f : \mathscr{M} \to \mathbb{F}$ defined by*

$$\mu_f(E) = \int_E f \, d\mu$$

is a signed measure.

In particular, if $\mathbb{F} = \mathbb{R}$ and $f = f^+ - f^-$ is the decomposition of f into is positive and negative parts, then

$$\mu_f = \mu_{f^+} - \mu_{f^-}.$$

Proof. First consider the case $\mathbb{F} = \mathbb{R}$. If f is unsigned, then we have already seen μ_f is a finite positive measure. Dropping the assumption that f is unsigned, consider the decomposition of f into is positive and negative parts, $f = f^+ - f^-$. Each of $\mu_{f^\pm}$ is a finite positive measure and hence so is $\mu_f = \mu_{f^+} - \mu_{f^-}$.

For the complex case, write $f = g + ih$, for real valued functions $h, g \in \mathcal{L}^1(\mu)$ and apply the already proven real case of to each of h and g. □

Remark 6.4. The measure μ_f is often denoted $f \, d\mu$.

Note that in the real case the real case, by choosing $E_+ = \{f \geq 0\}$ and $E_- = \{f < 0\}$, the proof actually shows there exists finite positive measures $\mu_\pm$ such that $\mu_f = \mu_+ - \mu_-$ and also such that $\mu_\pm(E) = \mu_f(E \cap E_\pm)$ and of course $E_+ \cap E_- = \varnothing$. In particular, if $F \subseteq E_\pm$, the $\mu_\pm(F) \geq 0$. □

6.2. **Total variation.**

Definition 6.5. The *total variation*, $|\rho|$, of an $\mathbb{F}$-measure ρ on a measure space $(X, \mathscr{M})$ is the function $|\rho| : \mathscr{M} \to [0, \infty]$ defined for $E \in \mathscr{M}$ by

$$|\rho|(E) = \sup\{\sum_{j=1}^{n} |\rho(E_j)| : n \in \mathbb{N},\ \{E_1, \ldots, E_n\} \text{ is a measurable partition of } E\}.$$

Remark 6.6. In the notation of Definition 6.5, for $E \in \mathscr{M}$,

(i) $|\rho(E)| \leq |\rho|(E)$;

(ii) $\rho(E) = |\rho(E)|$ if ρ is a positive measure;

(iii) $|\rho(E)| = 0$ if and only if $\rho(A \cap E) = 0$ for all $A \in \mathscr{M}$;

(iv) If $F \in \mathscr{M}$ and $F \subseteq E$, then $|\rho|(F) \leq |\rho|(E)$;

(v) $|\rho|(E) = \sup\{\sum_{j=1}^{n} |\rho(E_j)| : E_1, \ldots, E_n \in \mathscr{M}$ are disjoint and $\cup_{j=1}^{n} E_j \subseteq E\}$.

(vi) If μ is a (positive) measure on $(X, \mathscr{M})$ and $|\rho(f)| \leq \mu(F)$ for all $F \in \mathscr{M}$, then $|\rho| \leq \mu$ in the sense that $|\rho|(E) \leq \mu(E)$ for all $E \in \mathscr{M}$.

(vii) If $\mathbb{F} = \mathbb{R}$, then $|\rho(E)| = \sup\{|\rho(A)| + |\rho(B)| : A, B \in \mathscr{M},\ A \cap B = \varnothing,\ A \cup B \subseteq E\}$.

Proposition 6.7. *Suppose $(X, \mathscr{M}, \mu)$ is a measure space. If $h \in \mathcal{L}^1(\mu)$, then $|h\,d\mu| = |h|\,d\mu$; that is $|h\,d\mu|(E) = \int_E |h|\,d\mu$ for all $E \in \mathscr{M}$.*

Before proving the theorem, we establish a couple of lemmas.

Lemma 6.8. *In the context of Proposition 6.7, $|h\,d\mu| \leq |h|\,d\mu$.*

Proof. For notational ease, let $\mu_h = h\,d\mu$. Given $E \in \mathscr{M}$ and a measurable partition $\{E_1, \ldots, E_n\}$ of E,

$$\sum_{j=1}^{n} |\mu_h(E_j)| = \sum_{j=1}^{n} \Big| \int_{E_j} h\,d\mu \Big| \leq \sum_{j=1}^{n} \int_{E_j} |h|\,d\mu = \sum_{j=1}^{n} \mu_{|h|}(E_j).$$

Thus $|h\,d\mu|(E) \leq |h|\,d\mu(E)$ as desired. □

END Wednesday 2025-03-05

Lemma 6.9. *Proposition 6.7 holds for measurable simple functions h.*

Proof. Since h is simple, there exists a measuarable partition $\{F_1, \ldots, F_m\}$ of X and scalars $c_1, \ldots, c_m \in \mathbb{F}$ such that $h = \sum_{k=1}^{m} c_k \chi_{F_k}$. (Since $h \in \mathcal{L}^1(\mu)$, for each k either $\mu(F_k) < \infty$ or $c_k = 0$.) Let $E \in \mathscr{M}$ and a measurable partition $\{E_1, \ldots, E_n\}$ of E be given. Thus,

$$h\chi_E = \sum_{j,k=1}^{m,n} c_k \chi_{E_j \cap F_k}$$

and $\{E_j \cap F_k : 1 \le j \le n,\ 1 \le k \le m\}$ is a measurable partition of E. Further,

$$\begin{aligned}\sum_{j,k=1}^{m,n} |\mu_h(E_j \cap F_k)| &= \sum_{j,k=1}^{m,n} |\int_{E_j \cap F_k} h\, d\mu| = \sum_{j,k=1}^{m,n} |c_k|\mu(E_j \cap F_k) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{m} \int_{E_j \cap F_k} |h|\, d\mu = \sum_{j=1}^{n} \int_{E_j} |h|\, d\mu \\ &= \int_E |h|\, d\mu = \mu_{|h|}(E).\end{aligned}$$

Thus $|h\, d\mu|(E) \ge |h|\, d\mu(E)$. An application of Lemma 6.8 completes the proof. □

Proof of Proposition 6.7. Let $\epsilon > 0$ be given. Since simple functions are dense in $\mathcal{L}^1(\mu)$, there is a simple function $g \in \mathcal{L}^1(\mu)$ such that $\|h - g\|_1 < \epsilon$. For a measurable set E,

$$\mu_{|g|}(E) = \int_E |g|\, d\mu \ge \int_E |h|\, d\mu - \int_E |h - g|\, d\mu \ge \mu_{|h|}(E) - \|h - g\|_1 > \mu_{|h|}(E) - \epsilon.$$

Thus $|g|\, d\mu(E) > |h|\, d\mu(E) - \epsilon$.

Similarly, for a measurable partition $\{E_1, \dots, E_n\}$ of E,

$$\begin{aligned}\sum_{j=1}^{n} |\mu_h(E_j)| = \sum_{j=1}^{n} |\int_{E_j} h\, d\mu| &\ge \sum_{j=1}^{n} |\int_{E_j} g\, d\mu| - \sum_{j=1}^{n} |\int_{E_j} (g - h)\, d\mu| \\ &\ge \sum_{j=1}^{n} |\mu_g(E_j)| - \sum_{j=1}^{n} \int_{E_j} |g - h|\, d\mu \\ &\ge \sum_{j=1}^{n} |\mu_g(E_j)| - \|g - h\|_1 \ge \sum_{j=1}^{n} |\mu_g(E_j)| - \epsilon.\end{aligned}$$

Thus, $|h\, d\mu|(E) \ge |g\, d\mu|(E) = \epsilon$ and using Lemma 6.9, $|h\, d\mu|(E) \ge |g\, d\mu|(E) - \epsilon = |g|\, d\mu(E) - \epsilon > |h|\, d\mu(E) - 2\epsilon$. It follows that $|h\, d\mu|(E) \ge |h|\, d\mu(E)$. □

Proposition 6.10. *If ρ is an $\mathbb{F}$-measure on a measurable space $(X, \mathscr{M})$, then $|\rho|$ is a (positive) measure on $\mathscr{M}$.*

Later we will see that $|\rho|$ is a finite measure.

Proof. Since $\rho(\varnothing) = 0$ it follows that $|\rho|(\varnothing) = 0$.

Now suppose $E_1, E_2, \dots$ is a disjoint sequence from $\mathscr{M}$ and let $E = \cup_{k=1}^{\infty} E_j$. Given measurable partitions $\{E_{k,1}, \dots, E_{k,n_k}\}$ of E_k, for each $N \in \mathbb{N}$ the collection of sets $\{E_{k,j} : 1 \le k \le N,\ 1 \le j \le n_k\}$ is a finite disjoint collection of measurable sets such that $\cup_{k=1}^{N} \cup_{j=1}^{n_k} E_{k,j} \subseteq E$. Thus, using Remark 6.6 item v,

$$\sum_{k=1}^{N} \sum_{j=1}^{n_k} |\rho(E_{k,j})| \le |\rho|(E).$$

Since $\sum_{j=1}^{n_k} |\rho(E_{k,j})| \le |\rho|(E_k)$ for each k, it follows that

$$\sum_{k=1}^{N} |\rho|(E_k) \le |\rho|(E)$$

and therefore

$$\sum_{k=1}^{\infty} |\rho|(E_k) \le |\rho|(E).$$

To prove the reverse inequality, let $\{F_1, \dots, F_n\}$ be a given measurable partition of E. Thus, $E_{j,k} = F_j \cap E_k$ for $1 \le j \le n$ are disjoint measurable subsets E_k. By Remark 6.6 item (v), $\sum_{j=1}^{n} |\rho(F_j \cap E_k)| \le |\rho|(E_k)$ for each k. Therefore, since ρ is an $\mathbb{F}$-measure,

$$\begin{aligned}\sum_{j=1}^{n} |\rho(F_j)| &= \sum_{j=1}^{n} |\sum_{k=1}^{\infty} \rho(F_j \cap E_k)| \le \sum_{j=1}^{n}\sum_{k=1}^{\infty} |\rho(F_j \cap E_k)| \\ &= \sum_{k=1}^{\infty}\sum_{j=1}^{n} |\rho(F_j \cap E_k)| \le \sum_{k=1}^{\infty} |\rho|(E_k).\end{aligned}$$

Thus $|\rho|(E) \le \sum_{k=1}^{\infty} |\rho|(E_k)$ and the proof is complete. □

Proposition 6.11. *If ρ is an $\mathbb{F}$ measure on the measurable space $(X, \mathscr{M})$, then $|\rho|$ is a finite measure. Equivalently, $|\rho|(X) < \infty$.*

The proof of Proposition 6.11 uses the following lemma whose proof, based upon disjointification, is the same as for the case of a finite positive measure.

Lemma 6.12. *If ρ is an $\mathbb{F}$-measure on the measurable space $(X, \mathscr{M})$ and $E_0 \supseteq E_1 \supseteq E_2 \dots$ is a decreasing sequence of measurable sets, then*

$$\rho(\cap_{n=1}^{\infty} E_n) = \lim_{n\to\infty} \rho(E_n).$$

Similarly, if $E_0 \subseteq E_1 \subseteq E_2 \dots$ is an increasing sequence from $\mathscr{M}$, then

$$\rho(\cup_{n=1}^{\infty} E_n) = \lim \rho(E_n).$$

END Friday 2025-03-07

Proof of Proposition 6.11. As a first step in proving the result in the case that ρ is a real measure, suppose $E \in \mathscr{M}$ and $|\rho|(E) = \infty$ and let $C > 0$ be given. By item (vii) of Remark 6.6, there exists disjoint sets $A, B \in \mathscr{M}$ such that $A, B \subseteq E$ and $|\rho(A)| + |\rho(B)| \ge 2(C + |\rho(E)|)$. Thus, without loss of generality,

$$|\rho(A)| \ge C + |\rho(E)|.$$

Since $|\rho|$ is a measure (Theorem 6.11),

$$|\rho|(A) + |\rho|(E \setminus A)| = |\rho|(E) = \infty,$$

so that either $|\rho|(A) = \infty$ or $|\rho|(E \setminus A)| = \infty$. We will show that $|\rho(A)|$, $|\rho(E \setminus A)| \geq C$, which shows that E contains a subset $F \in \mathscr{M}$ such that $|\rho|(F) = \infty$ and $|\rho(F)| \geq C$.

By construction, $|\rho(A)| \geq C$. From $\rho(E \setminus A) + \rho(A) = \rho(E)$, it follows that

$$|\rho(E \setminus A)| = \big|\rho(E) - \rho(A)\big| = \big|\rho(A) - \rho(E)\big| \geq |\rho(A)| - |\rho(E)| \geq C.$$

To prove the proposition still assuming ρ is a real measure, it suffices to show $|\rho|(X) < \infty$. Arguing by contradiction, suppose $|\rho|(X) = \infty$. Choosing $A = X$ and $C = 1$, there is a measurable set $E_1 \subseteq X$ such that $|\rho(E_1)| \geq 1$ and $|\rho|(E_1) = \infty$. Suppose now measurable sets $E_1 \supseteq E_2 \subseteq E_3 \cdots \supseteq E_m$ have been constructed such that $|\rho(E_n)| \geq n$ and $|\rho|(E_n) = \infty$ for $1 \leq n \leq m$. It follows that, with $A = E_m$ and $C = m+1$, there is a measurable set $E_{m+1} \subseteq E_m$ such that $|\rho(E_{m+1})| \geq m+1$ and $|\rho|(E_{m+1}) = \infty$. Thus recursion produces a nested decreasing sequence of measurable sets $(E_n)_{n=1}^{\infty}$ such that $|\mu(E_n)| \geq n$. An application of Lemma 6.12 produces the contradiction that $|\rho(\cap E_n)| = \infty$.

To complete the proof, suppose now ρ is a $\mathbb{F}$-measure. From what is already proved, $|\operatorname{real}\rho|(X), |\operatorname{image}\rho|(X) < \infty$. On the other hand, $|\rho(F)| \leq |\operatorname{real}\rho(F)| + |\operatorname{image}\rho(F)|$. Hence $|\rho|(E) \leq |\operatorname{real}\rho|(E) + |\operatorname{image}\rho|(E)$ and thus $|\rho|(X)|\operatorname{real}\rho|(X) + |\operatorname{image}\rho|(X) < \infty$. □

6.3. **Banach spaces of measures.**

Proposition 6.13. *Suppose $(X, \mathscr{M})$ is a measurable space. The mapping $\|\cdot\| : \mathbb{M}_{\mathbb{F}}(\mathscr{M}) \to [0,\infty)$ defined by $\|\rho\| = |\rho|(X)$ is a norm on the space of measures.*

Proof. Suppose $\rho, \tau \in \mathbb{M}_{\mathbb{F}}(\mathscr{M})$ and $c \in \mathbb{F}$. From Remark 6.6 item (vii), $\|\rho\| = 0$ if and only if $\rho = 0$. It is straightforward to verify that $\|c\rho\| = |c|\,\|\rho\|$.

Finally, to prove the triangle inequality, simply note that

$$\|\rho + \tau\| = |(\rho+\tau)(X)| = |\rho(X) + \tau(X)| \leq |\rho(X)| + |\tau(X)| = \|\rho\| + \|\tau\|. \qquad \square$$

Proposition 6.14. *Suppose $(X, \mathscr{M})$ is a measurable space. The normed space $\mathbb{M}_{\mathbb{F}}(\mathscr{M})$ is a Banach space.*

It is straightforward to show that if $(\rho_n)_n$ is a Cauchy sequence from $\mathbb{M}_{\mathbb{F}}(\mathscr{M})$, then, for each $E \in \mathscr{M}$ the sequence $(\rho_n(E))_n$ converges. That $\rho : \mathscr{M} \to \mathbb{F}$ given by $\rho(E) = \lim_n \rho_n(E)$ that $\rho \in \mathbb{M}_{\mathbb{F}}(\mathscr{M})$ and (ρ_n) converges to ρ (in the normed vector space $\mathbb{M}_{\mathbb{F}}(\mathscr{M})$) is left to the gentle reader.

6.4. **The Hahn decomposition.**

Definition 6.15. Suppose ρ is an $\mathbb{R}$-measure on a measurable space $(X, \mathscr{M})$. A set $E \in \mathscr{M}$ is *totally positive* (resp. totally negative) for ρ if $\rho(F \cap E) \geq 0$ (resp. $\rho(F \cap E) \leq 0$) for all $F \in \mathscr{M}$; the set E *totally null* if $\rho(F \cap E) = 0$ for all $F \in \mathscr{M}$.

Remark 6.16. A set E is totally null for ρ if and only if it is both totally positive and totally negative for ρ.

A set $E \in \mathscr{M}$ is totally positive for ρ if and only if $\rho(F) \leq \rho(E)$ for all $F \subseteq E$. (Consider $E \setminus F$).

If E is totally positive for ρ, then $\widetilde{\rho} : \mathscr{M} \to \mathbb{F}$ defined by $\widetilde{\rho}(F) = \rho(E \cap F)$ is a finite positive measure.

If $(E_n)_n$ is a sequence of totally positive sets, then $\cup_{n=1}^{\infty} E_n$ is also totally positive; that is X_+ is totally positive for ρ_f and X_- is totally negative for ρ_f. □

Example 6.17. In the context of Proposition 6.7, decompose a real-valued function $f \in \mathcal{L}^1(\mu)$ into its positive and negative parts $f = f^+ - f^-$, the sets $X_+ := \{x : f^+(x) > 0\}$ and $X_- := \{x : f^-(x) > 0\}$ are disjoint and totally positive for μ_f.

Theorem 6.18 (Hahn Decomposition Theorem)**.** *If ρ is an $\mathbb{R}$-measure on the measurable space $(X, \mathscr{M})$, then there exists a partition of X into disjoint measurable totally positive sets $X = X_+ \cup X_-$.*

The decomposition is unique in the sense that if X'_+, X'_- is another such pair, then $X_+ \Delta X'_+$ and $X_- \Delta X'_-$ are totally null for ρ.

END Monday 2025-03-10

The following lemma will be used in the proof of Theorem 6.18

Lemma 6.19. *Suppose ρ is an $\mathbb{R}$-measure on a measurable space $(X, \mathscr{M})$. If $\rho(G) > 0$, then there exists a subset $E \subseteq G$ such that E is totally positive and $\rho(E) \geq \rho(G)$.*

The proof uses the *greedy algorithm*.

Proof. For notational convenience, let $E_1 = G$.

If E_1 is totally positive, then there is nothing to prove. Otherwise, there is a measurable set $H \subseteq E_1$ such that $\rho(H) < 0$ and thus $\rho(E_1 \setminus H) > \rho(E_1)$; that is, there is a measurable set $F \subseteq E_1$ such that $\rho(F) > \rho(E_1)$. Thus the set

$$J_1 = \{n \in \mathbb{N}^+ : \text{ there is an } F \subseteq E_1 \text{ such that } \rho(F) \geq \rho(E_1) + \frac{1}{n}\}$$

is nonempty and therefore has a smallest element n_1. Choose E_2 such that $\rho(E_2) \geq \rho(E_1) + 1/n_1$. If E_2 is totally positive, then the proof is complete. Otherwise, let n_2 denote the smallest element of

$$J_2 = \{n \in \mathbb{N}^+ : \text{ there is an } F \subseteq E_2 \text{ such that } \rho(F) \geq \rho(E_2) + \frac{1}{n}\}$$

and choose $E_3 \subseteq E_2$ such that $\rho(E_3) \geq \rho(E_2) + 1/n_2$. (Note that $n_2 \geq n_1$ and it could be the case that $n_2 = n_1$.) Either this recursion terminates after finitely many steps producing a totally positive subset E of E_1 with $\rho(E) \geq \rho(E_1)$; or it generates a nested decreasing sequence of measurable sets (E_j) and a sequence of positive integers (n_j) such that

$$\rho(E_{j+1}) \geq \rho(E_j) + \frac{1}{n_j},$$

where

(39) $$n_j = \min\{n \in \mathbb{N}^+ : \text{ there is an } F \subseteq E_j \text{ such that } \rho(F) \geq \rho(E_j) + \frac{1}{n}\}.$$

In particular,

(40) $$\rho(E_{j+1}) \geq \sum_{k=1}^{j} \frac{1}{n_k} + \rho(E_1).$$

Assuming this latter case, let $X_+ = \bigcap_{j=1}^{\infty} E_j$. We will show that $\rho(X_+) > \rho(E_1)$ and X_+ is totally positive.

By Lemma 6.12, $\rho(E_j)$ increases to $\rho(E) > \rho(G)$. Thus, by equation (40) and the assumption that ρ is a finite measure, the sequence $(n_j)_j$ converges to infinity (as otherwise $\rho(E) = \infty$). To show that E must be totally positive, suppose, by way of contradiction, there exists a measurable $F \subseteq E$ such that $\rho(F) > \rho(E)$. There is an $m \in \mathbb{N}^+$ such that $\rho(F) > \rho(E) + 1/m$. There is a j such that $n_j > m$. Now $F \subseteq E_j$ and

$$\rho(F) > \rho(E) + 1/m > \rho(E_j) + \frac{1}{m} > \rho(E_j) + 1/m,$$

contradicting the choice of n_j in equation (39) and completing the proof. □

Proof of Theorem 6.18. If $\rho(G) \leq 0$ for all $G \in \mathscr{M}$, then ρ is totally negative and the choices $X_+ = \varnothing$ and $X_- = X$ satisfies the conclusion of the theorem.

Otherwise, by Proposition 6.11,

$$\infty > a = \sup\{\rho(G) : G \in \mathscr{M}\} > 0.$$

For each $n \in \mathbb{N}$ such that $a > 1/n$, there exists a set G_n such that $\rho(G_n) > a - 1/n > 0$. By Lemma 6.19, there exists a totally positive set E_n such that $E_n \subseteq G_n$ and $\rho(E_n) \geq \rho(G_n) > 0$. Let $F_m = \cup_{n=1}^{m} E_n$ and note that F_m is totally positive by Remark 6.16. Thus $\rho(F_m \setminus E_{m+1}) \geq 0$ and therefore

$$\rho(F_{m+1}) = \rho(E_{m+1}) + \rho(F_m \setminus E_{m+1}) \geq \rho(E_{m+1}).$$

By Lemma 6.12, $(\rho(F_m))_m$ converges to $\rho(E)$ where $E = \cup^{\infty} F_m = \cup^{\infty} E_n$. On the other hand, $(\rho(F_m))$ converges to a. Hence $\rho(E) = a$. By Remark 6.16, E is totally positive. Further, if $F \subseteq E^c$, then $a \geq \rho(E \cup F) = \rho(E) + \rho(F) = a + \rho(F)$ and therefore $\rho(F) \leq 0$. Hence E^c is totally negative and $\{E, E^c\}$ partition X.

For the final statement, observe if $F \subseteq X_+ \setminus X'_+ = X_+ \cap X'_-$, then $\rho(F) \geq 0$ since F is a subset of the totally positive set X_+; also $\rho(F) \leq 0$ since F is a subset of the totally negative set X'_-. Hence $\rho(F) = 0$ and thus $X_+ \setminus X'_+$ is totally null. The remaining details are left to the gentle reader. □

6.5. **The Jordan decomposition.**

Definition 6.20. Suppose ρ is an $\mathbb{F}$-measure on a measurable space $(X, \mathscr{M})$. A set E is a *support set* for ρ if E^c is totally null for ρ. Two signed measures ρ, σ are *mutually singular*, denoted $\rho \perp \sigma$, if they have disjoint support sets; that is, there exists disjoint measurable sets E and F such that E^c is totally null for ρ and F^c is totally null for σ.

Remark 6.21. Two positive measures ρ and σ on the same measurable space $(X, \mathscr{M})$ are mutually singular if and only if there exists disjoint (measurable) sets E and F such that $\rho(E^c) = 0 = \sigma(F^c)$ (in which case it can be assumed that $F = E^c$ if desired).

Example 6.22. Let m denote Lebesgue measure on $(\mathbb{R}, \mathcal{L})$ (where $\mathcal{L}$ is the sigma-algebra of Lebesgue measurable subsets of $\mathbb{R}$) and let $\delta : \mathcal{L} \to \mathbb{R}$ denote point mass at 0; that is

$$\delta(E) = \begin{cases} 1 & \text{if } 0 \in E \\ 0 & \text{if } 0 \notin E. \end{cases}$$

It is immediate that $m \perp \delta$.

Theorem 6.23 (Jordan Decomposition)**.** *If ρ is an $\mathbb{R}$-measure on $(X, \mathscr{M})$, then there exist unique positive measures ρ_+, ρ_- such that $\rho_+ \perp \rho_-$ and $\rho = \rho_+ - \rho_-$. Moreover, $|\rho| = \rho_+ + \rho_-$.*

Proof. Let $X = X_+ \cup X_-$ be a Hahn decomposition for ρ (Theorem 6.18) and define $\rho_\pm : \mathscr{M} \to \mathbb{R}$ by $\rho_\pm(E) = \pm\rho(E \cap X_\pm)$. It is immediate from the properties of the Hahn decomposition that ρ_+, ρ_- have the desired properties; uniqueness is left as an exercise.

To prove the last statement let $\tau = \rho_+ + \rho_1$. For $E \in \mathscr{M}$,

$$|\rho|(E) \geq |\rho(E \cap X_+)| + |\rho(E \cap X_-)| = \rho_+(E) + \rho_-(E) = \tau(E).$$

On the other hand,

$$\begin{aligned} \tau(F) = \rho_+(F) + \rho_-(F) &= |\rho(F \cap X_+)| + |\rho(F \cap X_-)| \\ &\geq |\rho(F \cap X_+) + \rho(F \cap X_-)| = |\rho(F)| \end{aligned}$$

for $F \in \mathscr{M}$ and hence $\tau \geq |\rho|$ by Remark 6.6 item (vi). □

END Wednesday 2025-03-12

Example 6.24. Referring to Proposition 6.3 and example 6.17, it is now immediate that the decomposition $\mu_f = \mu_{f^+} - \mu_{f^-}$ is the Jordan decomposition of μ_f. Thus the Jordan decomposition theorem is analogous to the decomposition of a real-valued function into its positive and negative parts.

6.6. **The Radon-Nikodym derivative.**

Definition 6.25. Suppose $(X,\mathscr{M},\mu)$ is a measure space. An $\mathbb{F}$-measure $\rho:\mathscr{M}\to\mathbb{F}$ is *absolutely* continuous with respect to μ, written $\rho\ll\mu$ provided $\rho(E)=0$ whenever $E\in\mathscr{M}$ and $\mu(E)=0$.

Remark 6.26. Given a measure space $(X,\mathscr{M},\mu)$ and an $f\in\mathcal{L}^1(\mu)$, the measure $\mu_f:\mathscr{M}\to\mathbb{F}$ defined by

$$\mu_f(E)=\int_E f\,d\mu$$

(see Proposition 6.3) is absolutely continuous with respect to μ. That is $\mu_f=f\,d\mu\ll\mu$.

If ρ is an $\mathbb{F}$-measure on $(X,\mathscr{M})$, then $\rho\ll|\rho|$. □

Theorem 6.27 (Radon-Nikodym)**.** *Suppose μ and ν are σ-finite positive measures on a measurable space $(X,\mathscr{M})$. If $\nu\ll\mu$, then there exists (an essentially unique) measurable function $h:X\to[0,\infty)$ such that $\nu=\mu_f$; that is*

$$\nu(E)=\int_E h\,d\mu$$

for all $E\in\mathscr{M}$.

In the case that ν is finite, $h\in L^1(\mu)$.

The function h is the *Radon-Nikodym derivative* of ν with respect to μ, denoted $\frac{d\nu}{d\mu}$.

Corollary 6.28. *Suppose $(X,\mathscr{M},\mu)$ is a σ-finite measure space and $\rho:\mathscr{M}\to\mathbb{F}$ is an $\mathbb{F}$-measure. If $\rho\ll\mu$, then there exists an $h\in L^1(\mu)$ such that $\rho=h\,d\mu$.*

Proof. Suppose ρ is an $\mathbb{R}$-measure and let $\rho=\rho_+-\rho_-$ denote its Jordan decomposition. It is routine to check if $\mu(E)=0$, then both $\rho_\pm(E)=0$. Two applications of Theorem 6.27 produces unsigned functions $h_\pm\in L^1(\mu)$ such that $\rho_\pm=h_\pm\,d\mu$. Hence the function $h=h_+-h_-$ is in $L^1(\mu)$ and $\rho=h\,d\mu$.

Finally, if ρ is a $\mathbb{C}$-measure, then $\operatorname{real}\rho$ and $\operatorname{image}\rho$ are $\mathbb{R}$-measures. Thus there exists $h,g\in L^1(\mu)$ such that $\operatorname{real}\rho=h\,d\mu$ and $\operatorname{image}\rho=g\,d\mu$. Hence $(h+ig)\in L^1(\mu)$ and $\rho=(h+ig)\,d\mu$. □

Corollary 6.29. *If ρ is an $\mathbb{F}$-measure on a measurable space $(X,\mathscr{M})$, then, there exists an $h\in L^1(\mu)$ such that $|h|=1$ a.e. $|\rho|$ and $\rho=h\,d\mu$; that is*

$$\rho(E)=\int_E h\,d|\rho|$$

for all $E\in\mathscr{M}$.

Proof. Note that ρ is absolutely continuous with respect to $\mu=|\rho|$ and hence the Radon-Nikodym Theorem, Theorem 6.27, produces an $h\in L^1(\mu)$ such that $\rho=h\,d\mu$. To see that $|h| = 1$ almost everywhere μ, note that Proposition 6.7 implies $\mu = |\rho| = |h|\,d\mu$. In particular,

$$\int_E (1 - |h|)\, d\mu = 0$$

and the result follows by choosing $E = \{|h| \neq 1\}$. □

A consequence of the Lebesgue-Radon-Nikodym theorem is the existence of *conditional expectations*.

Corollary 6.30. *Suppose $(X, \mathscr{M}, \mu)$ be a σ-finite measure space (μ a positive measure), $\mathscr{N}$ a sub-σ-algebra of $\mathscr{M}$, and $\nu = \mu|_{\mathscr{N}}$ is σ-finite. If $f \in L^1(\mu)$ then there exists $g \in L^1(\nu)$ (unique modulo ν-null sets) such that*

$$\int_E f\, d\mu = \int_E g\, d\nu$$

for all $E \in \mathscr{N}$. (The function g is called the conditional expectation *of f on $\mathscr{N}$.)*

Sketch of proof. Since f is $\mathscr{M}$-measurable, it is also $\mathscr{N}$ measurable and moreover $f \in L^1(\nu)$ Thus, we may define $\rho : \mathscr{N} \to \mathbb{F}$ by

$$\rho(E) = \int_E f\, d\nu = \int_E f\, d\mu. \tag{41}$$

It is immediate that ρ is absolutely continuous with respect to ν. Thus, by Corollary 6.28, there is an essentially unique $g \in L^1(\nu)$ such that equation 41 holds. □

END Friday 2025-03-14

Proof of Theorem 6.27. As a first step, assume μ and ν are finite positive measures. This step is key and the proof given here is due to von Neumann.

Let $m = \nu + \mu$. Thus m is a finite positive measure and in particular $1 \in L^2(m)$, where $L^2(m)$ denotes the real Hilbert space of real valued square integrable (with respect to m) functions. Given $f \in L^2(m)$, observe that the Cauchy-Schwarz inequality gives

$$\int |f|\, d\nu \le \int |f|\, 1\, dm = \left|\langle |f|,\, 1\rangle_{L^2(m)}\right| \le \|f\|_2\, \|1\|_2. \tag{42}$$

Hence it is sensible to define $\varphi : L^2(m) \to \mathbb{R}$ by

$$\varphi(f) = \int f\, d\nu$$

and moreover, the estimate of equation (42) gives $|\varphi(f)| \le \|1\|_2\, \|f\|_2$ so that φ is a bounded linear functional (with norm at most $\|1\|_2$).

By the Reisz representation Theorem (for Hilbert space), Theorem 5.28, there exists a $g \in L^2(m)$ (real-valued) such that

$$\int f\, d\nu = \varphi(f) = \int f\, g\, dm = \int f\, g\, d\mu + \int f\, g\, d\nu.$$

and therefore,

$$\int f(1-g)\,d\nu = \int f\,g\,d\mu \tag{43}$$

for all $f \in L^2(\mu)$. With $G_1 = \{g \geq 1\}$, equation (43) we have

$$0 \leq \mu(G_1) = \int_{G_1} d\mu = \int \chi_{G_1} g\,d\mu = \int \chi_{G_1}(1-g)\,d\nu \leq 0.$$

Hence $\mu(G_1) = 0$. Since $\nu \ll \mu$, it also is the case the $\nu(G_1) = 0 = m(G_1)$.[5] Now let $G_n = \{g < -\frac{1}{n}\}$ and observe, again using equation (43) that

$$-\frac{1}{n}\mu(G_n) \geq \int \chi_{G_n} g\,d\mu = \int \chi_{G_n}(1-g)\,d\nu \geq 0.$$

Thus $\mu(G_n) = 0$ and hence $\mu(G_0) = 0$, where $G_0 = \cup_{G_n} = \{g < 0\}$. As before it follows that $\nu(G_0) = 0 = m(G_0)$.

Since $m(G_0) = 0 = m(G_1)$, it is harmless to assume, as we now do, that $0 \leq g < 1$ pointwise. Let $\psi = \frac{1}{1-g}$, set $h = g\psi$ and note both ψ and h are unsigned. The sequence (ψ_n) defined by

$$\psi_n = \psi\chi_{\{\psi \leq n\}}$$

is a pointwise increasing sequence of bounded non-negative functions that converges pointwise to ψ. Thus each $\psi_n \in L^2(m)$ and by the monotone convergence theorem (twice) and equation (43) the sequence

$$\int \psi_n(1-g)\chi_E\,d\nu = \int \psi_n\,g\,\chi_E\,d\mu$$

converges to both $\int \chi_E\,d\nu$ and $\int h\chi_E\,d\mu$; that is

$$\nu(E) = \int_E h\,d\mu = \mu_h(E)$$

for all $E \in \mathscr{M}$. Choosing $E = X$ shows $h \in L^1(\mu)$ completing the proof in this special case that both ν and μ are finite.

We now sketch a proof of the case that both ν and μ are σ-finite (positive) measures. Since ν and μ are σ-finite, there exists a sequence $X_1 \subseteq X_2 \subseteq \ldots$ of measurable sets of finite measure such that $X = \cup X_n$. For $n \in \mathbb{N}$ define $\mu_n : \mathscr{M} \to [0,\infty)$ by $\mu_n(E) = \mu(E \cap X_n)$ and define ν_n similarly. The pair (ν_n, μ_n) are finite positive measures and $\nu_n \ll \mu_n$. Hence, by what has already been proved and with $n = 1$, there exists $h_1 \in L^1(\mu_1)$ such that $d\nu_1 = h_1\,d\mu_1$. Without loss of generality, we assume $h_1 = 0$ on X_n^c. With $n = 2$, there is an $h_2 \in L^2(\mu)$ such that $d\nu_2 = h_2\,d\mu_2$ and $h_2 = 0$ on X_2^c. Moreover, since h_1 and h_2 agree μ a.e. on X_1, we also assume, without loss of generality, that $h_1 = h_2$ on X_1. Continuing in this fashion constructs an increasing sequence (h_n) of unsigned functions that converges pointwise to some h and satisfies $\nu_n = h_n\,d\mu_n$ for each n. Finally, given $E \in \mathscr{M}$, let $E_n = E \cap E_n$ and apply the monotone convergence theorem

[5] Compare with the proof of Theorem 6.31 in Subsection 6.7, where absolute continuity is not assumed.

to the sequences χ_{E_n} and $h\chi_{E_n}$ and the measures ν and μ respectively to conclude that the sequence

$$\int \chi_{E_n}\, d\nu = \nu_n(E) = h_n d\mu_n(E) = \int h\chi_{E_n} d\mu$$

converges to both $\nu(E)$ and to $\int_E h\, d\mu$. Hence $\nu = h\, d\mu$. In the case that ν is finite, choosing $E = X$ and using $\nu(X) < \infty$ gives $h \in L^1(\mu)$. □

6.7. **The Lebesgue decomposition.**

Theorem 6.31 (Lebesgue Decomposition - positive measure version)**.** *Suppose $(X, \mathscr{M}, \mu)$ is a σ-finite measure space. If ν is a finite positive measure on $(X, \mathscr{M})$, then there exist unique positive measures ν_a and ν_s such that*

(i) $\nu_a \ll \mu$;
(ii) $\nu_s \perp \mu$; *and*
(iii) $\nu = \nu_a + \nu_s$.

Moreover, there exists a measurable set F such that

(i) $\nu_a(E) = \nu(E \cap F) = 0$;
(ii) $\nu_s(E) = \nu(E \cap F^c) = 0$; *and*
(iii) $\mu(F^c) = 0$.

Remark 6.32. The result holds if ν is assumed σ-finite, a result that follows easily from the case of ν finite. The details are left to the interested reader. □

The uniqueness asserted in Theorem 6.31 is a consequence of the following lemma.

Lemma 6.33. *Suppose $(X, \mathscr{M}, \mu)$ is a measure space and $\rho : \mathscr{M} \to \mathbb{R}$ is a measure. If $\rho \ll \mu$ and $\rho \perp \mu$, then $\rho = 0$.*

Proof. Since $\rho \perp \mu$, there exist a set $F \in \mathscr{M}$ such that $\rho(E) = \rho(E \cap F^c)$ and $\mu(E) = \mu(E \cap F)$ for all $E \in \mathscr{M}$. Hence, for $E \in \mathscr{M}$,

$$\mu(E \cap F^c) = \mu(E \cap F^c \cap F) = 0.$$

Since $\rho \ll \mu$, it follows that $\rho(E \cap F^c) = 0$. Thus both F and F^c are totally null for ρ. Hence $\rho = 0$. □

Proof of Theorem 6.31. Let $m = \nu + \mu$. In particular, m is σ-finite and $\nu \ll m$. Hence, by the Radon Nikodym Theorem, Theorem 6.27. there is a uniquely (a.e. m) determined unsigned function $g : X \to [0, \infty)$ such that $\nu = m_g$. Thus, for all measurable E,

$$\nu(E) = \int_E g\, d(\mu + \nu) = \int_E g\, dm$$

An easy argument shows $g \leq 1$ a.e. m.

Let $F = \{g < 1\}$ and note $F^c = \{g = 1\}$.[6] Define $\nu_a(E) = \nu(E \cap F)$ and $\nu_s(E) = \nu(E \cap F^c)$ for $E \in \mathscr{M}$. Both are positive measures, $\nu_s(F) = 0$ and $\nu = \nu_a + \nu_s$. Next,

$$\nu(F^c) = \int_{F^c} g \, d(\mu + \nu) = \mu(F^c) + \nu(F^c).$$

Since $\nu(F^c) \in [0, \infty)$ it follows that, $\mu(F^c) = 0$ and thus $\nu_s \perp \mu$.

To prove ν_a is absolutely continuous with respect to μ, suppose E is measurable and $\mu(E) = 0$. Letting $F_n = \{g < 1 - \frac{1}{n}\} \subseteq F$ (for positive integers n),

$$\nu(E \cap F_n) \le (1 - \frac{1}{n})[\mu + \nu](E \cap F_n) = (1 - \frac{1}{n})\nu(E \cap F_n).$$

Hence $\nu(E \cap F_n) = 0$. Since $E \cap F = \cup(E \cap F_n)$ it follows that $\nu_a(E) = \nu(E \cap F) = 0$ and therefore $\nu_a \ll \mu$.

To prove uniqueness, suppose $\nu = \rho_a + \rho_s$. Since these are finite measures, $\rho_a - \nu_a = \nu_s - \rho_s$. Now the $\mathbb{R}$-measure on the right hand side is singular with respect to μ while the $\mathbb{R}$-meaure on the left hand side is absolutely continuous with respect to μ. Hence, by Lemma 6.33, both are 0. □

END Monday 2025-03-24

The Lebesgue decomposition easily extends to the case of $\mathbb{F}$-measures

Theorem 6.34 (Lebesgue Decomposition - $\mathbb{F}$-measure version)**.** *Suppose $(X, \mathscr{M}, \mu)$ is a σ-finite measure space. If $\rho : \mathscr{M} \to \mathbb{F}$ is an $\mathbb{F}$-measure, then then there exist unique measures ρ_a and ρ_s such that $\rho_a \ll \mu$ and $\rho_s \perp \mu$ and $\rho = \rho_a + \rho_s$.*

Proof. Let $\nu = |\rho|$. By Theorem 6.31 there exists measures ν_s and ν_a such that $\nu_a \perp \mu$ and $\nu_s \ll \mu$. Moreover, there exists a measurable set F such that $\nu_a(E) = \nu(E \cap F)$ and $\nu_s(E) = \nu(E \cap F^c)$ for $E \in \mathscr{M}$ and $\mu(F^c) = 0$. In particular, $\nu_s(F) = 0$. Define $\rho_a, \rho_s : \mathscr{M} \to \mathbb{F}$ by $\rho_a(E) = \rho(E \cap F)$ and $\rho_s(E) = \rho(E \cap F^c)$. If $E \subseteq F$ is measurable, then $\rho_s(E) = \rho(E \cap F^c) = \rho(\varnothing) = 0$ and hence F is totally null for ρ_s. Thus $\rho_s \perp \mu$, since also $\mu(F^c) = 0$. On the other hand, if $\mu(E) = 0$, then, writing $E = (E \cap F) \cup (E \cap F^c)$

$$|\rho_a(E)| = |\rho(E \cap F)| \le |\rho|(E \cap F) = \nu_a(E) = 0,$$

since $\nu_a \ll \mu$. Hence $\rho_a(E) = 0$ and we conclude that $\rho_a \ll \mu$. By construction $\rho = \rho_a + \rho_s$.

Once again, uniqueness follows from Lemma 6.33. □

6.8. **Duality for Lebesgue spaces - conclusion.** This subsection contains a sketch of a proof, based on the Radon-Nikodym Theorem (Theorem 6.27), that the isometric map of Proposition 4.14 is in fact onto (unitary). Recall, given $1 \le p \le \infty$ and a $g \in L^q(\mu)$, where q is the conjugate index to p, that for $f \in L^p(\mu)$, Hölder's inequality

[6] In the case that $\nu \ll \mu \ll m$ the set F^c is m-null. Compare with the proof of Theorem 6.27.

(Theorem 4.8) implies $gf \in L^1(\mu)$ and moroever $\|fg\|_1 \le \|f\|_p \|g\|_q$. Thus, we obtain a bounded linear functional $L_g : L^p(\mu) \to \mathbb{F}$ of norm at most $\|g\|_q$ defined by

$$L_g(f) = gf.$$

Let $\Phi : L^q(\mu) \to L^p(\mu)^*$ denote the bounded map (with norm at most one) given by $\Phi(g) = L_g$.

Theorem 6.35. *If $(X, \mathscr{M}, \mu)$ is a σ-finite measure space and $1 \le p < \infty$, then the mapping $\Phi : L^q(\mu) \to L^p(\mu)^*$ defined by $\Phi(g) = L_g$ is an isometric isomorphism.*

Recall, Problems 2.12 and 4.6 says that the result fails in the case of $p = \infty$. Likewise the result fails for $p = 1$ without the σ-finite hypothesis. See Remark 4.15.

Proof of Theorem 6.35 in the case of a finite measure. Proposition 4.14 says Φ is isometric. Thus it remains to show that Φ is onto under the assumption that μ is a finite (positive) measure. Let $\varphi \in L^p(\mu)^*$ be given. Define $\nu : \mathscr{M} \to \mathbb{F}$ as follows. Given $E \in \mathscr{M}$ the function $\chi_E \in L^1(\mu)$ since μ is finite. Set $\nu(E) = \varphi(\chi_E)$. In particular, $\nu(\varnothing) = 0$. To prove that ν is countably additive and hence an $\mathbb{F}$-measure, suppose $(E_n)_{n=1}^\infty$ is a sequence of disjoint measurable sets and let $E = \cup_{j=1}^\infty E_j$. Let $s_n = \sum_{j=1}^n \mathbf{1}_{E_j}$. In particular, (s_n) increases pointwise with limit $s = \mathbf{1}_E$. Further, $0 \le (s - s_n)^p \le 1 \in L^p(\mu)$ and thus, by dominated convergence, (s_n) converges to s in $L^p(\mu)$.[7] Using continuity and linearity of φ,

$$\nu(E) = \varphi(s) = \lim_{n\to\infty} \varphi(s_n) = \lim_{n\to\infty} \sum_{j=1}^n \nu(E_j) = \sum_{j=1}^\infty \nu(E_j).$$

If $\mu(E) = 0$, then $\chi_E = 0$ in L^p and therefore $\nu(E) = 0$. Thus, the measure ν is absolutely continuous with respect to μ. Consequently, by the Radon-Nikodym Theorem, there exists an $h \in L^1(\mu)$ such that

$$\varphi(\chi_E) = \nu(E) = \int_E h\, d\mu. \tag{44}$$

Temporarily, view L_h as defined (and continuous) on $L^\infty(\mu)$. (See Proposition 4.14.)

If s is a measurable simple function, then, by equation (44),

$$\varphi(s) = \int_X s\, h\, d\mu.$$

Now suppose f is a bounded unsigned measurable function. Since μ is a finite measure, f is in $L^p(\mu)$ as well as $L^\infty(\mu)$. Hence, both $\lambda_h(f)$ and $\varphi(f)$ are defined. There exists a sequence (s_n) of measurable simple functions $0 \le s_n \le f$ such that (s_n) converges to f uniformly and therefore in both $L^p(\mu)$ and $L^\infty(\mu)$ (again because μ is finite). It follows

$$\varphi(f) = \lim \varphi(s_n) = L_h(s_n) = L_h(f).$$

[7] On the other hand, (s_n) does not necessarily converge to s in $L^\infty(\mu)$.

It now follows that if f is bounded and measurable, then $\varphi(f) = L_h(f)$.

To prove $h \in L^q$, first assume $p > 1$. For positive integers N, let $E_N = \{|h| \leq N\}$ and let $h_N = h\chi_{E_n}$. Thus h_N is bounded and so is $f_N = \overline{h_N}|h_N|^{q-2}$ (where we set $f_N(x) = 0$ if $x \notin E_N$). Thus all are in each $L^r(\mu)$ since μ is finite. By $L^p(\mu)$ continuity of φ,

$$\|f_N\|_p \, \|\varphi\| \geq |\varphi(f_N)| = |L_h(f_N)| = \int_X f_N h \, d\mu = \int_X |h_N|^q \, d\mu = \|h_N\|_q^q. \tag{45}$$

On the other hand, $\|f_N\|_p = \|h_N\|_q^{q-1}$, and combining this equality with (45) we see that $\|h_N\|_q \leq \|\varphi\|$. By monotone congvergence, $h \in L^q(\mu)$ and moreover $\|h\|_q \leq \|\varphi\|$.

In the $p = 1$ case, put $E_t = \{|h| > t\}$ and let $f_t = \frac{\overline{h}}{|h|}\chi_{E_t}$. Thus $\|f_t\|_1 = \mu(E_t)$ for all t, and, since $f \in L^\infty(\mu)$,

$$\mu(E_t) \, \|\varphi\| = \|\varphi\| \, \|f_t\|_1 \geq |\varphi(f_t)| = |L_h(f_t)| = \int_X f_t \, h \, d\mu = \int_{E_t} |h| \, d\mu \geq t \, \mu(E_t). \tag{46}$$

Hence $\mu(E_t) = 0$ for $t > \|\varphi\|$ and thus $h \in L^\infty(\mu)$ and in fact $\|\varphi\| \geq \|h\|_\infty$.

Now that we know $h \in L^q$, it follows that L_h is continuous. It also agrees with φ on simple functions. Since simple functions are dense in $L^q(\mu)$, the conclusion $\varphi = L_h$ follows. □

END Wednesday 2025-03-26

The following lemmas will be used to prove Theorem 6.35 in the case μ is σ-finite.

Lemma 6.36. *If μ $(X, \mathscr{M}, \mu)$ is a σ-finite measure space, then there exists a measurable function $w \in L^1(\mu)$ such that $0 < w(x) < 1$ for all x.*

Proof. Write $X = \cup_{n=1}^\infty X_n$, a countable union of disjoint measurable sets of finite measure. Let $w_n = \frac{1}{2^n(\mu(X_n)+1)}\mathbf{1}_{X_n}$ and $w = \sum_{n=1}^\infty w_n$. □

Lemma 6.37. *Suppose μ is a σ-finite measure, $w \in L^1(\mu)$ and $0 < w(x) < 1$ for all x. Let τ denote the measure $w \, d\mu$.*

For $1 < p < \infty$, a measurable function f is in $L^p(\mu)$ if and only if $g = w^{-\frac{1}{p}} f \in L^p(\tau)$ and in this case $\|f\|_p = \|g\|_p$; that is, the mapping $\Psi_p : L^p(\tau) \to L^p(\mu)$ defined by

$$\Psi_p(f) = w^{\frac{1}{p}} f$$

is a (linear) isometric isomorphism.

Proof. It is easy to check that Ψ_p is isometric with inverse given by $f \mapsto w^{-\frac{1}{p}} f$. □

Proof of Theorem 6.35. For the σ-finite case with $p > 1$, let w be as in Lemma 6.36. Likewise, let $\tau = w \, d\mu$. By Lemma 6.37 the mappings $\Psi_r : L^p(\tau) \to L^p(\mu)$ defined by $\Psi_r h = w^{\frac{1}{r}} h$ are linear isometric isomorphisms. Thus, $\psi = \varphi \circ \Phi_p$ is a bounded linear functional on $L^p(\tau)$. Since τ is a finite measure, by what is already proved, there is a

$g \in L^q(\tau)$ such that $\psi = L_g$. Let $h = \Psi_q g = w^{\frac{1}{q}} h$. Thus $h \in L^q(\mu)$ and $\|h\|_q = \|g\|_q$. Moreover, if $f \in L^p(\mu)$, then $F := \Psi_p^{-1} f = w^{-\frac{1}{p}} f \in L^p(\tau)$ and

$$\begin{aligned}\varphi(f) =&\psi(F) = L_g(F)\\ =&\int F\, g\, d\tau = \int F\, g\, w\, d\mu\\ =&\int w^{\frac{1}{p}} F\,(w^{\frac{1}{q}} g)\, d\mu = \int f\, h\, d\mu = L_h(f).\end{aligned}$$

In the case $p = 1$, write $X = \cup_{n=1}^{\infty} X_n$, where X_n are measurable sets of finite measure and apply what has already been proven in the to the measure space $(X_n, \mathscr{M}_n, \mu_n)$, where $\mathscr{M}_n = \{E \in \mathscr{M} : E \subseteq X_n\}$ and $\mu_n = \mu|_{\mathscr{M}_n}$. The details are left to the gentle reader. □

6.9. **Problems.**

Problem 6.1. a) Prove Proposition 6.30. b) In the case μ = Lebesgue measure on $[0, 1)$, fix a positive integer k and let $\mathscr{N}$ be the sub-σ-algebra generated by the intervals $[\frac{j}{k}, \frac{j+1}{k})$ for $j = 0, \dots k-1$. Give an explicit formula for the conditional expectation g in terms of f. c) Show that the σ-finite hypothesis on ν is needed.

7. The Fourier transform

We assume all functions are complex-valued unless stated otherwise.

Definition 7.1. [The Fourier transform] Let $f \in L^1(\mathbb{R})$. The *Fourier transform* of f is the function $\mathcal{F}(f) = \widehat{f} : \mathbb{R} \to \mathbb{C}$ defined at each $t \in \mathbb{R}$ by

$$\widehat{f}(t) := \int_{-\infty}^{\infty} f(x) e^{-2\pi i t x}\, dx. \tag{47}$$

The terminology *Fourier transform* is often used for the mapping that sends f to $\widehat{f}$. □

Note that $\widehat{f}$ makes sense, since $f \in L^1(\mathbb{R})$ and, for each t, the function $\exp(-2\pi i t x) \in L^\infty(\mathbb{R})$. In fact, $|\widehat{f}(t)| \le \|f\|_1$ so that $\widehat{f} \in F_b(\mathbb{R})$, the Banach space of bounded functions on $\mathbb{R}$ with the supremum norm, and $\|\hat{f}\|_\infty \le \|f\|_1$.

The basic properties of the Fourier transform listed in the following proposition stem from two basic facts: (1) Lebesgue measure is translation invariant; and (2) that, for each $t \in \mathbb{R}$, the function

$$\chi_t : x \to \exp(2\pi i t x)$$

is a *character* of the additive group $(\mathbb{R}, +)$;[8] that is, χ_t is a homomorphism from $\mathbb{R}$ into the mulitplicative group of unimodular complex numbers. Explicitly for all $x, y, t \in \mathbb{R}$

$$\chi_t(x + y) = \chi_t(x)\chi_t(y).$$

[8]The characters of the multiplicative group $\mathbb{T}$ (the unit circle in the complex plane) are parameterized by $\mathbb{Z}$ with $n \in \mathbb{Z}$ corresponding to the character $\chi_n(\gamma) = \gamma^n$; that is $\chi_n(e^{it}) = e^{int}$ (for $t \in \mathbb{R}$). Proceeding in this way lead to the theory of Fourier Series.

Example 7.2. Given real numbers $a < b$, let $f = \chi_{[a,b]}$ (that the interval is closed, open, or neither is not important here) and verify

$$\widehat{f}(t) = \begin{cases} i\frac{e^{-2\pi ibt}-e^{-2\pi iat}}{2\pi t} & t \neq 0 \\ b-a & t = 0. \end{cases}$$

Note that the derivative of $\exp(-2\pi i\alpha t)$ at $t = 0$ is $-2\pi i\alpha$. In particular, for $b > 0$ and $f = \frac{1}{2b}\chi_{[-b,b]}$,

$$\widehat{f}(t) = \operatorname{sinc}(2\pi bt).$$

Since for $k \in \mathbb{N}$,

$$\int_{k\pi}^{(k+1)\pi} \left|\frac{\sin(x)}{(k+1)}\right| \geq \frac{2}{k+1},$$

the sinc function is not in $L^1(\mathbb{R})$; that is, in general $f \in L^1(\mathbb{R})$, does not imply $\widehat{f} \in L^1(\mathbb{R})$. □

The following example will be used later when the Poisson kernel for the upper half plane is introduced.

Example 7.3. For $a > 0$, let

$$Q_a(t) = e^{-2a\pi|t|} \tag{48}$$

(the extra factor of 2π turns out to be a convenient normalization).

$$\widehat{Q_a}(-t) = \int_{\mathbb{R}} Q_a(t)e^{2\pi itx}\, dt = \frac{1}{\pi}\frac{a}{a^2+x^2}.$$

Note, this function Q is, in a sense, a smoother version of the indicator function from example 7.2. It has the virtue that is Fourier transform is in $L^1(\mathbb{R})$. □

7.1. **Basic Properties.** Before going further we introduce some notation: for fixed $y \in \mathbb{R}$ and a function $f : \mathbb{R} \to \mathbb{C}$, define $f_y(x) := f(x-y)$.

Proposition 7.4 (Basic properties of the Fourier transform)**.** *Let $f, g \in L^1(\mathbb{R})$ and let $\alpha \in \mathbb{R}$.*

(a) (Linearity) $\widehat{cf+g} = c\widehat{f} + \widehat{g}$
(b) (Translation) $\widehat{f_y}(t) = e^{-2\pi ity}\widehat{f}(t)$
(c) (Modulation) If $g(x) = e^{2\pi i\alpha x}f(x)$, then $\widehat{g}(t) = \widehat{f}(t-\alpha)$
(d) (Reflection) If $g(x) = \overline{f(-x)}$, then $\widehat{g}(t) = \overline{\widehat{f}(t)}$.
(e) (Scaling) If $\lambda > 0$ and $g(x) = f(x/\lambda)$ then $\widehat{g}(t) = \lambda\widehat{f}(\lambda t)$.

Proof. Each of these properties is verified by elementary transformations of the integral defining $\widehat{f}$; the details are left as an exercise. □

Proposition 7.5. *If $f \in L^1(\mathbb{R})$, then $\widehat{f}$ is continuous and bounded ($\widehat{f} \in C_b(\mathbb{R})$) and $\|\widehat{f}\|_\infty \leq \|f\|_1$. In particular, the mapping $L^1(\mathbb{R}) \ni f \mapsto \widehat{f} \in C_b(\mathbb{R})$ is a bounded linear map of norm at most* 1.

Proof. Fix $t \in \mathbb{R}$ and a sequence $t_n \to t$. The sequence $f(x)e^{-2\pi i t_n x}$ converges to $f(x)e^{-2\pi i t x}$ pointwise on $\mathbb{R}$, and since trivially $|f(x)e^{-2\pi i t_n x}| \leq |f(x)|$ for all n, we have by dominated convergence

$$\begin{aligned}
\widehat{f}(t) &= \int_{-\infty}^{\infty} f(x)e^{-2\pi i t x}\,dx \\
&= \int_{-\infty}^{\infty} \lim_{n\to\infty} [f(x)e^{-2\pi i t_n x}]\,dx \\
&= \lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)e^{-2\pi i t_n x}\,dx \\
&= \lim_{n\to\infty} \widehat{f}(t_n).
\end{aligned}$$

The second statement of the theorem follows immediately from the estimate $\sup_{t\in\mathbb{R}} |\widehat{f}(t)| \leq \|f\|_1$. □

In fact, $\widehat{f}$ always belongs to $C_0(\mathbb{R})$, a result that is known as the *Riemann-Lebesgue Lemma*. To prove it we first need the following result, which we will apply often (recall the notation $f_y(x) := f(x-y)$):

Lemma 7.6 (Translation is continuous on L^p)**.** *If $1 \leq p < \infty$ and $f \in L^p(\mathbb{R})$, then $\lim_{y\to 0} \|f_y - f\|_p = 0$. In particular, if (y_n) converges to y, then $(f_{y_n})_n$ converges to f_y in $L^p(\mathbb{R})$.*

Sketch. We sketch the proof of an approximation argument, leaving the details as an exercise. Let $X \subseteq L^p$ denote the set of all f for which the conclusion of the theorem is true.

Verify X is a vector space and contains χ_I for all finite intervals I. By Littlewood's first principle, a measurable set of finite measure is nearly a finite union of intervals. Thus X contains the indicator functions of all sets of finite measure and therefore all simple L^p functions. Since simple L^p functions are dense in L^p, it suffices to show that X is closed. Toward this end, note if $f, g \in L^p$ and $\|f-g\|_p < \epsilon$, then $\|f-g\|_p = \|f_y - g_y\|_p < \epsilon$ for all $y \in \mathbb{R}$ by the translation invariance of Lebesgue measure. Now suppose that g is in the closure of X and let $\epsilon > 0$ be given. Choose $f \in X$ with $\|f-g\|_p < \epsilon$, and choose $\delta > 0$ so that $\|f_y - f\|_p < \epsilon$ for all $|y| < \delta$. Then for all $|y| < \delta$,

$$\|g_y - g\|_p < \|g_y - f_y\|_p + \|f_y - f\|_p + \|f-g\|_p < 3\epsilon.$$

Thus $g \in X$ as well and hence X is closed. The proof is finished. □

END Friday 2025-03-28

Lemma 7.7 (The Riemann-Lebesgue Lemma)**.** *If* $f \in L^1(\mathbb{R})$*, then* $\widehat{f} \in C_0(\mathbb{R})$*.*

Proof. From Proposition 7.5, $\widehat{f}$ is continuous.

The proof here that f vanishes at infinity appeals to the continuity of translation in L^1 (Lemma 7.6), and a simple trick. First, since $e^{-\pi i} = -1$,

$$\widehat{f}(t) = -\int_{\mathbb{R}} f(x) e^{-2\pi i t(x+(1/2t))}\,dx = -\int_{\mathbb{R}} f\left(x - \frac{1}{2t}\right) e^{-2\pi i x t}\,dx. \tag{49}$$

Combining equation (49) with the usual definition of $\widehat{f}$, we have

$$\widehat{f}(t) = \frac{1}{2}\int_{\mathbb{R}} \left(f(x) - f\left(x - \frac{1}{2t}\right)\right) e^{-2\pi i x t}\,dx.$$

Thus

$$|\widehat{f}(t)| \leq \frac{1}{2}\|f - f_{\frac{1}{2t}}\|_1.$$

By Lemma 7.6 $\|f - f_{\frac{1}{2t}}\|_1 \to 0$ as $t \to \pm\infty$. □

Continuing our catalog of basic properties, we see that the Fourier transform also interacts nicely with differentiation.

Proposition 7.8 (Multiplication becomes differentiation)**.** *Suppose* $f \in L^1(\mathbb{R})$*. If* $g(x) := x f(x)$ *belongs to* $L^1(\mathbb{R})$*, then* $\widehat{f}$ *is differentiable and*

$$\widehat{g}(t) = \frac{-1}{2\pi i}\frac{d}{dt}\widehat{f}(t),$$

for all $t \in \mathbb{R}$*.*

The proof uses the standard estimate,

$$|1 - e^{it}| \leq |t|$$

for t real and dominated convergence.

Proof. For real numbers $s \neq t$

$$\frac{\widehat{f}(s) - \widehat{f}(t)}{s - t} = \int_{-\infty}^{\infty} \frac{e^{-2\pi i s x} - e^{-2\pi i t x}}{s - t} f(x)\,dx. \tag{50}$$

The estimate

$$\left|\frac{e^{-2\pi i s x} - e^{-2\pi i t x}}{s - t}\right| \leq 2\pi|x|$$

holds for all $s \neq t$. Thus, by the assumption $xf(x) \in L^1$, a dominated convergence argument in (50) shows that the limit as $s \to t$ exists and moreover

$$\begin{aligned}\lim_{s\to t}\frac{\widehat{f}(s)-\widehat{f}(t)}{s-t} &= \lim_{s\to t}\int_{-\infty}^{\infty}\frac{e^{-2\pi isx}-e^{-2\pi itx}}{s-t}f(x)\,dx\\ &= \int_{\infty}^{\infty}(-2\pi i)e^{-2\pi itx}xf(x)\,dx\\ &= -2\pi i\widehat{g}(t).\end{aligned}$$

Thus $\widehat{f}$ is differentiable and the claimed formula holds. □

Note that if $f \in L^1$ and also $g(x) := x^n f(x) \in L^1$ for some integer $n \geq 1$, then $x^k f(x)$ belongs to L^1 for all $0 \leq k \leq n$. The previous proposition can then be applied inductively to conclude:

Corollary 7.9. *If $f \in L^1$ and $g := x^n f \in L^1$, then $\widehat{f}$ is n times differentiable, and*

$$\widehat{x^k f} = \left(\frac{-1}{2\pi i}\right)^k \widehat{f}^{(k)} \quad \textit{for each } 0 \leq k \leq n.$$

One also expects a theorem in the opposite direction: the Fourier transform should convert differentiation to multiplication by the independent variable. Under reasonable hypotheses, this is the case.

Proposition 7.10. *If $f \in C_0(\mathbb{R})$ and f' is continuous and in L^1, then*

$$\mathcal{F}(f')(t) = \widehat{f'}(t) = 2\pi it\hat{f}(t).$$

Proof. Compute

$$\begin{aligned}\widehat{f'}(t) &= \int_{-\infty}^{\infty} f'(x)e^{-2\pi itx}\,dx\\ &= \lim_{b\to\infty}\int_{-b}^{b} f'(x)e^{-2\pi itx}\,dx\\ &= \lim_{b\to\infty}\left([f(b)e^{-2\pi ibt}-f(-b)e^{2\pi ibt}]+2\pi it\int_{-b}^{b}f(x)e^{-2\pi ixt}\,dx\right)\\ &= 2\pi it\widehat{f}(t),\end{aligned}$$

where the second equality follows from the Dominated Convergence Theorem, the third using integration by parts, and the fourth from the $C_0(\mathbb{R})$ assumption on f and another application of Dominated Convergence. □

Example 7.11. For $a > 0$ let $g = g_a$ denote the *Gaussian*,

$$g(x) := e^{-\pi ax^2}.$$

(The factor of π will be convenient given our choice of normalization in the definition of the Fourier transform.)

Rather than computing the transform of g_a directly, we exploit Propositions 7.8 and 7.10. We may also assume $a = 1$ since the general case follows from this by scaling (Proposition 7.4(e)). Note that $h \in L^1(\mathbb{R})$ and $g' = -2\pi h$. Thus,

$$\begin{aligned}(\widehat{g})'(t) &= -2\pi i \widehat{h}(t) \\ &= -2\pi i \mathcal{F}(-\frac{1}{2\pi} g') \\ &= i 2\pi i t \hat{g}(t) \\ &= -2\pi t \hat{g}(t),\end{aligned} \tag{51}$$

where the first equality follows from Proposition 7.8, the second from $g' = -2\pi h$ and the third from Proposition 7.10. It follows from equation (51) and the product rule that

$$\frac{d}{dt}(e^{\pi t^2} \widehat{g}(t)) = 0.$$

Hence the function $e^{\pi t^2} \widehat{g}(t)$ is constant. To evaluate the constant, we set $t = 0$ and use the well-known Gaussian integral

$$\widehat{g}(0) = \int_{-\infty}^{\infty} e^{-\pi x^2} \, dx = 1.$$

We note in passing that $\mathcal{F}(h_1) = -i h_1$ too.

As a final remark, the $\mathcal{F}(H_{4n} g_1) = H_{4n} g_1$, where H_n are (appropriately normalized) hermite polynomials. $\square$

7.2. **Convolution and the Fourier transform.** The last set of basic properties of the Fourier transform concern its interaction with convolution, which we now introduce. If f, g are measurable functions on $\mathbb{R}$, the *convolution* of f and g is the function

$$(f * g)(x) := \int_{\mathbb{R}} f(x-y) g(y) \, dy \tag{52}$$

defined at each x for which the integral makes sense. In particular, if $f \in L^\infty$ and $g \in L^1$, then $f * g$ is defined on all of $\mathbb{R}$. Observe, using the invariance of Lebesgue measure with respect to $x \to -x$ and a simple change of variable,

$$f * g(x) = \int_{\mathbb{R}} g(x-y) f(y) \, dy = g * f(x). \tag{53}$$

END Monday 2025-03-31

Index