MAA6617 COURSE NOTES
SPRING 2020

Contents

20. Normed vector spaces

Let $\mathcal{X}$ be a vector space over a field $\mathbb{K}$ (in this course we always have either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$).

Definition 20.1. A *norm* on $\mathcal{X}$ is a function $\|\cdot\| : \mathcal{X} \to \mathbb{R}$ satisfying:

(i) (positivity) $\|x\| \geq 0$ for all $x \in \mathcal{X}$, and $\|x\| = 0$ if and only if $x = 0$;
(ii) (homogeneity) $\|kx\| = |k|\|x\|$ for all $x \in \mathcal{X}$ and $k \in \mathbb{K}$, and
(iii) (triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathcal{X}$.

$\triangleleft$

Using these three properties it is straightforward to check that the quantity

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

defines a metric on $\mathcal{X}$. The resulting topology is the *norm topology.* The next proposition is simple but fundamental; it says that the norm and the vector space operations are continuous in the norm topology.

Proposition 20.2 (Continuity of vector space operations)**.** *Let $\mathcal{X}$ be a normed vector space over $\mathbb{K}$.*

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 $\mathbb{K}$ 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. $\square$

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

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \quad \text{for all } x \in \mathcal{X}.$$

Equivalent norms determine the same topology on $\mathcal{X}$ and the same Cauchy sequences (Problem 20.2). A normed space is a *Banach space* if it is complete in the norm topology. 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.

The following proposition gives a convenient criterion for a normed vector space to be complete. A series $\sum_{n=1}^{\infty} x_n$ in $\mathcal{X}$ is *absolutely convergent* if $\sum_{n=1}^{\infty} \|x_n\| < \infty$. The series *converges* in $\mathcal{X}$ if the limit $\lim_{N\to\infty} \sum_{n=1}^{N} x_n$ exists in $\mathcal{X}$ (in the norm topology). (Quite explicitly, the series $\sum_{n=1}^{\infty} x_n$ converges to $x \in \mathcal{X}$ if $\lim_{N\to\infty} \left\| x - \sum_{n=1}^{N} x_n \right\| = 0$.)

Proposition 20.3. *A normed space $(\mathcal{X}, \|\cdot\|)$ is complete if and only if every absolutely convergent series in $\mathcal{X}$ is convergent.*

Proof. 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

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

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

Conversely, suppose every absolutely convergent series in $\mathcal{X}$ is convergent. Given a Cauchy sequence (x_n) from X, choose a super-Cauchy subsequence (y_k); i.e., $(y_k = x_{n_k})_k$ and

$$\sum_{k=1}^{\infty} \|y_{k+1} - y_k\| < \infty.$$

(To do this, first choose N_1 such that $\|x_n - x_m\| < 2^{-1}$ for all $n, m \geq N_1$. Next choose $N_2 > N_1$ such that $\|x_n - x_m\| < 2^{-2}$ for all $n, m \geq N_2$. Continuing in this way recursively defines an increasing sequence of integers $(N_k)_{k=1}^{\infty}$ such that $\|x_n - x_m\| < 2^{-k}$ for all $n, m \geq N_k$. Set $y_k = x_{N_k}$.) The series $\sum_{k=1}^{\infty} (y_{k+1} - y_k)$ is absolutely convergent and hence, by hypothesis, convergent in $\mathcal{X}$. In other words, the sequence $(y_k - y_1)_k$ of partial sums converges in $\mathcal{X}$ which means that (x_n) has a convergent subsequence. The proof is finished by invoking a standard fact about convergence in metric spaces: if (x_n) is a Cauchy sequence which has a convergent subsequence, then the full sequence converges. □

20.1. **Examples.**

(a) Of course, $\mathbb{K}^n$ 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 $\mathbb{K}^n$ 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 $1 \leq p < \infty$, and the ℓ^∞*-norm*

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

For $1 \leq p < \infty$ and $p \neq 2$, 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 $(1 \leq p \leq \infty)$ are equivalent on $\mathbb{K}^n$ (though the constants c, C depend on the dimension n). 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 20.3.

(b) (Sequence spaces) Define

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

It is a simple exercise to check that each of these is a vector space (a subspace of the vector space of all functions $f : \mathbb{N} \to \mathbb{K}$). Define, for functions $f : \mathbb{N} \to \mathbb{K}$,

$$\|f\|_\infty := \sup_m |f(m)|$$
$$\|f\|_1 := \sum_{m=1}^{\infty} |f(m)|.$$

Then $\|f\|_\infty$ defines a norm on both c_0 and ℓ^∞, and $\|f\|_1$ is a norm on ℓ^1. Equipped with these respective norms, each is a Banach space. We sketch the proof for c_0. Verification of the other two assertions is left as exercises (Problem 20.4).

The key observation is that (f_n) converges to f in the $\|\cdot\|_\infty$ norm if and only if (f_n) converges to f uniformly as functions on $\mathbb{N}$. Suppose (f_n) is a Cauchy sequence in c_0. Then the sequence of functions f_n is uniformly Cauchy on $\mathbb{N}$, and in particular converges pointwise to a function f.

We must show that this f belongs to c_0 and that $f_n \to f$ uniformly on $\mathbb{N}$. For this, let $\epsilon > 0$ be given. There is an N so that $\|f_m - f_n\| < \epsilon$ for $m, n \geq N$. Thus, for each $n \geq N$, all $\ell \in \mathbb{N}$ and all $m \geq n \geq N$, $|f_m(\ell) - f_n(\ell)| < \epsilon$. Thus, fixing $n = N$ and letting m tend to infinity, we have $|f(\ell) - f_N(\ell)| \leq \epsilon$ for all ℓ. There is an M so that $|f_N(\ell)| < \epsilon$ for $\ell \geq M$. Hence, for such ℓ,

$$|f(\ell)| \leq |f_N(\ell)| + \epsilon < 2\epsilon.$$

Thus $f \in c_0$ and (f_n) converges to f in c_0.

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

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

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 20.9).

(c) (L^1 spaces) Let $(X, \mathscr{M}, m)$ be a measure space. The quantity

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

defines a norm on $L^1(m)$, 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. Note that ℓ^1 from the last example is a special case (what is the measure space?))

Proposition 20.4. *$L^1(m)$ is a Banach space.*

Proof. It suffices to verify the hypotheses of Proposition 20.3. If $\sum_{n=1}^\infty f_n$ is absolutely convergent (so that $\sum_{n=1}^\infty \|f_n\|_1 < \infty$), then

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

Thus the function $g := \sum_{n=1}^\infty |f_n|$ belongs to L^1 and is thus finite m-a.e. by Tonelli's theorem. In particular the sequence of partial sums $s_N = \sum_{n=1}^N f_n$ is a sequence of measurable functions with $|s_N| \leq g$ 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. $\square$

(d) (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 $p > 1$). The sequence space ℓ^p is defined analogously: it is the set of $f : \mathbb{N} \to \mathbb{K}$ for which

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

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

When $p = \infty$, and the measure m is σ-finite, we define $L^\infty(m)$ to be the set of all functions $f : X \to \mathbb{K}$ with the following property: there exists $M > 0$ such that

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

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 $L^\infty(m)$ into a Banach space.

(e) ($C(X)$ spaces) Let X be a compact metric space and let $C(X)$ denote the set of continuous functions $f : X \to \mathbb{K}$. It is a standard fact from advanced calculus that the quantity $\|f\|_\infty := \sup_{x\in X} |f(x)|$ is a norm on $C(X)$. A sequence is Cauchy in this norm if and only if it is uniformly Cauchy. It is thus also a standard fact that $C(X)$ is complete in this norm—completeness just means that a uniformly Cauchy sequence of continuous functions on X converges uniformly to a continuous function.

This example can be generalized somewhat: let X be a locally compact metric space. Say a function $f : X \to \mathbb{K}$ *vanishes at infinity* if for every $\epsilon > 0$, there exists a compact set $K \subset X$ such that $\sup_{x \notin K} |f(x)| < \epsilon$. Let $C_0(X)$ denote the set of continuous functions $f : X \to \mathbb{K}$ that vanish at infinity. Then $C_0(X)$ is a vector space, the quantity $\|f\|_\infty := \sup_{x \in X} |f(x)|$ is a norm on $C_0(X)$, and $C_0(X)$ is complete in this norm. (Note that c_0 from above is a special case.)

(f) (Subspaces and direct sums) If $(\mathcal{X}, \|\cdot\|)$ is a normed vector space and $\mathcal{Y} \subset \mathcal{X}$ 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.)

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

$$\|(x, y)\|_\infty := \max(\|x\|_{\mathcal{X}}, \|y\|_{\mathcal{Y}}),$$
$$\|(x, y)\|_1 := \|x\|_{\mathcal{X}} + \|y\|_{\mathcal{Y}}$$

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

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

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

(g) (Quotient 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. If $\mathcal{M}$ is a *closed* subspace of $\mathcal{X}$, then the quantity

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

is a norm on $\mathcal{X}/\mathcal{M}$, called the *quotient norm*. (Geometrically, $\|x + \mathcal{M}\|$ is the distance in $\mathcal{X}$ from x to the closed set $\mathcal{M}$.) It turns out that if $\mathcal{X}$ is complete, so is $\mathcal{X}/\mathcal{M}$. See Problem 20.20.

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

20.2. **Linear transformations between normed spaces.**

Definition 20.5. 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 $C > 0$ such that $\|Tx\|_{\mathcal{Y}} \leq C\|x\|_{\mathcal{X}}$ for all $x \in \mathcal{X}$. $\triangleleft$

Remark 20.6. Note that in this definition it would suffice 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?) $\diamond$

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 20.7. *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,

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

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

$$\|Tx\| \leq \|T\|\,\|x\| \tag{2}$$

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

Proof. Suppose T is bounded so that there exists a $C > 0$ 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. By hypothesis, with $\epsilon = 1$ there exists $\delta > 0$ such that if $\|x\| < \delta$, then $\|Tx\| < 1$. Fix a nonzero vector $x \in \mathcal{X}$ and a real number $0 < \lambda < \delta$. The vector $\lambda x/\|x\|$ has norm less than δ, so

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

Rearranging this we find $\|Tx\| \leq (1/\lambda)\|x\|$ for all $x \neq 0$, which shows T is bounded; in fact we can take $C = \frac{1}{\delta}$.

The rest of the proof is left as an exercise. $\square$

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 called the *operator norm* of T.

Proposition 20.8. *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 $B(\mathcal{X},\mathcal{Y})$. For each $x\in\mathcal{X}$, we have

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \le \|T_n - T_m\|\|x\| \tag{3}$$

which shows that $(T_n x)$ is a Cauchy sequence in $\mathcal{Y}$. By hypothesis, $T_n x$ converges in $\mathcal{Y}$. Define $T:\mathcal{X}\to\mathcal{Y}$ by setting $Tx := y$. It is straightforward to check that T is linear.

Let B denote the closed unit ball in $\mathcal{X}$. The sequence $(T_n|_B)$ is uniformly Cauchy by equation (3) and converges pointwise to $T|_B$. Hence $T|_B$ is continuous and $(T_n|_B)$ converges uniformly to $T|_B$. It follows that T is continuous at 0 and therefore T is continuous. Since $\|T_n - T\| = \sup\{\|(T_n-T)x\| : x\in B\}$ and since $(T_n|_B)$ converges to $T|_B$ uniformly, it follows that (T_n) converges to T in $B(\mathcal{X},\mathcal{Y})$.

□

If $T\in B(\mathcal{X},\mathcal{Y})$ and $S\in B(\mathcal{Y},\mathcal{Z})$, then two applications of the the inequality (2) gives, for $x\in\mathcal{X}$,

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

and it follows that $ST\in B(\mathcal{X},\mathcal{Z})$ and $\|ST\|\le\|S\|\|T\|$. 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 20.9 (Extending bounded operators)**.** *Let $\mathcal{X}$, $\mathcal{Y}$ be normed vector spaces with $\mathcal{Y}$ complete, and $\mathcal{E}\subset\mathcal{X}$ 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 $\|\widetilde{T}\| = \|T\|$.*

Proof. Recall, if X, Y are metric spaces, Y is complete, $D\subset X$ 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 $\tilde{T}$ of T is in fact linear and $\|T\| = \|\tilde{T}\|$. Both are routine exercises. □

Remark: The completeness of $\mathcal{Y}$ is essential in the above proposition; Problem 20.11 suggests a counterexample.

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 $T:\mathcal{X}\to\mathcal{Y}$. 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.

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. An isometry need not be surjective, however. The normed vector spaces are *isometrically isomorphic* if there is an invertible isometry $T : \mathcal{X} \to \mathcal{Y}$.

20.3. **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 20.14

(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 as an operator (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.

(c) Consider c_{00} with the $\|\cdot\|_\infty$ norm. Let $a : \mathbb{N} \to \mathbb{K}$ be any function and define a linear transformation $T_a : c_{00} \to c_{00}$ by
$$T_a f(n) = a(n)f(n). \tag{4}$$
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, and one may check that the formula (4) 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.

20.4. **Problems.**

Problem 20.1. Prove Proposition 20.2.

Problem 20.2. Prove equivalent norms define the same topology and the same Cauchy sequences.

Problem 20.3. (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 20.2 to conclude that every finite-dimensional normed vector space is complete.

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

Problem 20.4. Finish the proofs from Example 20.1(b).

Problem 20.5. A function $f : [0,1] \to \mathbb{K}$ 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,

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

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 20.6. 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 20.7. 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 20.8. Prove, if $\mathcal{X}, \mathcal{Y}$ are normed spaces, then the operator norm is a norm on $B(\mathcal{X}, \mathcal{Y})$.

Problem 20.9. 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 $\|f - g\|_\infty \ge 1$ for all $g \in c_{00}$.)

Problem 20.10. 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 20.11. 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 20.9 can fail if $\mathcal{Y}$ is not complete.)

Problem 20.12. 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 20.13. 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 20.14. 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 $\{e_1, \dots, e_d\}$ 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 20.3.

Problem 20.15. Prove the claims in Example 20.3(c).

Problem 20.16. Let $g : \mathbb{R} \to \mathbb{K}$ 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 20.17. Prove the claims about direct sums in Example 20.1(f).

Problem 20.18. 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

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

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 $\|g - f\|_\infty \le \frac{1}{2} + \epsilon$ (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 20.19. Prove, if $\mathcal{X}$ is an infinite-dimensional normed space, then the unit ball $ball(\mathcal{X}) := \{x \in \mathcal{X} : \|x\| \le 1\}$ is not compact in the norm topology. (Hint: use the result of Problem 20.18 to construct inductively a sequence of vectors $x_n \in \mathcal{X}$ such that $\|x_n\| = 1$ for all n and $\|x_n - x_m\| \ge \frac{1}{2}$ for all $m < n$.)

Problem 20.20. (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 (see Example 20.1(g)).
b) Show that the quotient map $x \to x + \mathcal{M}$ has norm 1. (Use Problem 20.18.)
c) Prove, if $\mathcal{X}$ is complete, so is $\mathcal{X}/\mathcal{M}$.

Problem 20.21. 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 $\alpha \ne \beta$.)

21. 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.

Let $\mathcal{X}$ be a normed vector space over the field $\mathbb{K}$. A *linear functional* on $\mathcal{X}$ is a linear map $L : \mathcal{X} \to \mathbb{K}$. As one might expect, we are especially interested in bounded linear functionals. Since $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is complete, the vector space of bounded linear functionals $B(\mathcal{X}, \mathbb{K})$ is itself a Banach space (complete normed vector space). This space is called the *dual space* of $\mathcal{X}$ and is denoted $\mathcal{X}^*$. 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.

21.1. **Examples.** 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. If we fix $g \in \ell^1$, then the functional $L_g : c_0 \to \mathbb{K}$ defined by

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

is bounded, since

$$|L_g(f)| \le \sum_{n=0}^{\infty} |f(n)g(n)| \le \|f\|_\infty \sum_{n=0}^{\infty} |g(n)| = \|g\|_1 \|f\|_\infty.$$

This inequality shows that $\|L_g\| \leq \|g\|_1$. In fact, equality holds, and every bounded linear functional on c_0 is of this form:

Proposition 21.1. *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

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

and that $\|L_g\| \leq \|g\|_1$. 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 $g : \mathbb{N} \to \mathbb{K}$ 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 let $h \in c_{00}$ be the function

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

By definition $h \in c_{00}$ and $\|h\|_\infty \leq 1$. Note that $h = \sum_{n=0}^{N} h(n)e_n$. 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 $\|g\|_1 \leq \|L\|$. Moreover, the same calculation shows that $L = L_g$ when restricted to c_{00}, so by the uniqueness of extensions of bounded operators, $L = L_g$. Thus the map $g \to L_g$ is onto and

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

□

Proposition 21.2. *$(\ell^1)^*$ is isometrically isomorphic to ℓ^∞.*

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

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 21.12).

If $f \in L^1(m)$ and g is a bounded measurable function with $\sup_{x \in X} |g(x)| = M$, then the map

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

is a bounded linear functional of norm at most M. We will prove in Section **??** that the norm is in fact equal to M, and every bounded linear functional on $L^1(m)$ is of this type (at least when m is σ-finite).

If X is a compact metric space and μ is a finite, signed Borel measure on X, then

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

is a bounded linear functional on $C_{\mathbb{C}}(X)$ with norm $\|\mu\| = |\mu|(X)$ (see Problem 21.8). A version of the Riesz Markov Theorem says the converse is true too.

Theorem 21.3 (Riesz-Markov)**.** *Suppose X is a compact Hausdorff space. If $\lambda \in C(X)^*$, then there exists a unique regular Borel measure σ such that, for $f \in C(X)$,*

$$\lambda(f) = \int_X f\, d\sigma.$$

The result is true with both real and complex scalars. Focusing on the case of real scalars, the strategy is to write λ as the difference of positive linear functionals on $C(X)$ and apply the Riesz-Markov Theorem (twice). (For complex scalars, write λ in terms of its real and imaginary parts and apply the result in the real case (twice)).

A function $f \in C(X)$ is *positive* (really nonnegative) if $f(x) \geq 0$ for all $x \in X$, written $f \geq 0$. Let $C(X)^+$ denote the positive elements of $C(X)$. Given linear functionals $\lambda, \rho \in C(X)^*$, the inequality $\lambda \leq \rho$ means that $\lambda(f) \leq \rho(f)$ for all $f \in C(X)^+$.

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

Definition 21.4. Let $\mathcal{X}$ be a real vector space. A *Minkowski functional* is a function $p : \mathcal{X} \to \mathbb{R}$ such that $p(x + y) \leq p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for all $x, y \in \mathcal{X}$ and nonnegative $\lambda \in \mathbb{R}$. $\triangleleft$

For example, 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. More generally 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 21.5 (The Hahn-Banach Theorem, real version)**.** *Let $\mathcal{X}$ be a vector space over $\mathbb{R}$, p a Minkowski functional on $\mathcal{X}$, and $\mathcal{M}$ a subspace of $\mathcal{X}$. If L a linear functional on $\mathcal{M}$ such that $L(x) \leq p(x)$ 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) $L'(x) \leq p(x)$ for all $x \in \mathcal{X}$ (L' is dominated by p).

The proof will invoke *Zorns 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 is a relation that is reflexive, symmetric and transitive; that is, for all $x, y, z \in \mathcal{S}$

(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 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 $w \in S$ that has no successor; that is there does not exist an $s \in S$ such that $s \neq w$ and $w \preceq s$.

Theorem 21.6 (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.*

Proof of Theorem 21.5. The idea is to show that the extension can be done one dimension at a time and then infer the existence of an extension to the whole space by appeal to Zorn's lemma. We may of course assume $\mathcal{M} \neq \mathcal{X}$. So, fix a vector $x \in \mathcal{X} \setminus \mathcal{M}$ and consider the subspace $\mathcal{M} + \mathbb{R}x \subset \mathcal{X}$. For any $m_1, m_2 \in \mathcal{M}$, by hypothesis,

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

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

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

and thus

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

Now choose any real number λ satisfying

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

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

$$\begin{aligned} L(m) - \lambda &\leq p(m - x) \\ L(m) + \lambda &\leq p(m + x). \end{aligned} \tag{5}$$

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. We now check that $L'(y) \leq p(y)$ for all $y \in \mathcal{M} + \mathbb{R}x$. Accordingly, suppose $m \in \mathcal{M}$, $t \in \mathbb{R}$ and let $y = m + tx$. If $t = 0$ there is nothing to prove. If $t > 0$, then, in view of equation (5),

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

and a similar estimate shows that $L'(m + tx) \leq p(m + tx)$ for $t < 0$.

We have thus successfully extended L to $\mathcal{M} + \mathbb{R}x$. To finish the proof, 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 $L'(y) \leq p(y)$ on $\mathcal{N}$. Declare $(L'_1, \mathcal{N}_1) \preceq (L'_2, \mathcal{N}_2)$ if $\mathcal{N}_1 \subset \mathcal{N}_2$ and $L'_2|_{\mathcal{N}_1} = L'_1$. This relation $\preceq$ is a partial order on $\mathcal{L}$. An exercise shows, given any increasing chain $(L'_\alpha, \mathcal{N}_\alpha)$ in $\mathcal{L}$, it 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 $\preceq$. Since it always possible to extend to a strictly larger subspace, the maximal element must have $\mathcal{N} = \mathcal{X}$, and the proof is finished. □

The proof is a typical application of Zorn's lemma - one knows how to carry out a construction one step a time, but there is no clear way to do it all at once.

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

Corollary 21.7. *Suppose $\mathcal{X}$ is a normed vector space over $\mathbb{R}$, $\mathcal{M}$ is a subspace, and L is a bounded linear functional on $\mathcal{M}$. If $C \ge 0$ and $|L(x)| \le C\|x\|$ for all $x \in \mathcal{M}$, then there exists a bounded linear functional L' on $\mathcal{X}$ extending L such that $\|L'\| \le C$.*

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

Before obtaining further corollaries, we extend these results to the complex case. 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.

Proposition 21.8. *Let $\mathcal{X}$ be a vector space over $\mathbb{C}$. If $L : \mathcal{X} \to \mathbb{C}$ is a $\mathbb{C}$-linear functional, then $u(x) = ReL(x)$ defines an $\mathbb{R}$-linear functional on $\mathcal{X}$ and $L(x) = u(x) - iu(ix)$. Conversely, if $u : \mathcal{X} \to \mathbb{R}$ is $\mathbb{R}$-linear then $L(x) := u(x) - iu(ix)$ is $\mathbb{C}$-linear. If in addition $p : \mathcal{X} \to \mathbb{R}$ is a seminorm, then $|u(x)| \le p(x)$ for all $x \in \mathcal{X}$ if and only if $|L(x)| \le p(x)$ for all $x \in \mathcal{X}$.*

Proof. Problem 21.5.

□

Theorem 21.9 (The Hahn-Banach Theorem, complex version)**.** *Let $\mathcal{X}$ be a vector space over $\mathbb{C}$, 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 $|L(x)| \le p(x)$ 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 $x \in \mathcal{X}$.

Proof. The proof consists of applying the real Hahn-Banach theorem to extend the $\mathbb{R}$-linear functional $u = \mathrm{Re}L$ to a functional $u' : \mathcal{X} \to \mathbb{R}$ and then defining L' from u' as in Proposition 21.8. 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 21.10. *Let $\mathcal{X}$ be a normed vector space.*

(i) If $\mathcal{M} \subset \mathcal{X}$ is a subspace and $L : \mathcal{M} \to \mathbb{K}$ is a bounded linear functional, then there exists a bounded linear functional $L' : \mathcal{X} \to \mathbb{K}$ such that $L'|_{\mathscr{M}} = L$ 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 $x \neq y$ in $\mathcal{X}$, there exists $L \in \mathcal{X}^*$ such that $L(x) \neq L(y)$.*

(iv) (Linear functionals detect distance to subspaces) *If $\mathcal{M} \subset \mathcal{X}$ 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) $L(x) = \mathit{dist}(x, \mathcal{M}) = \inf_{y \in \mathcal{M}} \|x - y\| > 0$.*

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

(ii): Let $\mathcal{M}$ be the one-dimensional subspace of $\mathcal{X}$ spanned by x. Define a functional $L : \mathcal{M} \to \mathbb{K}$ 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 $\|L\| = 1$.

(iii): Apply (ii) to the vector $x - y$.

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

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

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

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 $T : \mathcal{Y} \to \mathcal{X}^*$.

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 $\hat{x} : \mathcal{X}^* \to \mathbb{K}$ 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 21.11. *(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 21.10(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 21.12. 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 21.6). Reflexive spaces often have nice properties. For instance, the distance from a point to a (closed) subspace is attained. On the other hand, by Propositions 21.1 and 21.2, c_0^{**} is isometrically isomorphic to ℓ^∞. In Problem 21.7 you will show that c_0 is not isometrically isomorphic to ℓ^∞ and so c_0 is not reflexive. 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 21.10 and 21.15.

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

Proposition 21.13 (Completion of normed spaces)**.** *If $\mathcal{X}$ is a normed vector space, then there is a Banach space $\overline{\mathcal{X}}$ and in isometric 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.

21.3. **Dual spaces and adjoint operators.** 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 $f : \mathcal{Y} \to \mathbb{K}$ is a linear functional, then $T^*f : \mathcal{X} \to \mathbb{K}$ defined by

$$(T^*f)(x) = f(Tx) \tag{6}$$

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 21.14. *Let $T : \mathcal{X} \to \mathcal{Y}$ be a bounded linear transformation. For $f \in \mathcal{Y}^*$, define T^*f by the formula (6). Then:*

i) T^*f belongs to $\mathcal{X}^*$, and T^* is a linear map from $\mathcal{Y}^*$ into $\mathcal{X}^*$.
ii) $T^* : \mathcal{Y}^* \to \mathcal{X}^*$ is bounded and $\|T^*\| = \|T\|$.

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

$$\|T^*f\| \leq \|f\|\|T\|. \tag{7}$$

Thus T^* maps $\mathcal{Y}^*$ into $\mathcal{X}^*$ and it is straightforward to verify that T^* is linear, which proves item (21.14). Moreover, the inequality of equation (7) 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 21.10(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, $\|T^*\| \geq \|T\|$. $\square$

21.4. **Duality for Sub and Quotient Spaces.** 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 20.20), 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 $\|\pi\| \leq 1$. Further, by Problem 20.18 (or see Proposition 21.15 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

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

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 21.15 (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 21.14 as does $\|\psi\| = \|\pi\| \leq 1$. To prove that ψ is isometric, let $\lambda \in (\mathcal{X}/\mathcal{M})^*$ be given. Automatically, $\|\psi(\lambda)\| \leq \|\lambda\|$. 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,

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

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

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

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

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 $\psi(\lambda) = \tau$. $\square$

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 21.10, 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 21.16 (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 $\|f\| \ge \|\lambda\|$. Hence,

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

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

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 $\tilde{T} : \mathcal{X}/\mathcal{M} \to \mathcal{Y}$. A canonical choice is $\mathcal{M} = \ker(T)$ in which case $\tilde{T}$ 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 21.17. *If $\mathcal{X}$ is a Banach space, $\mathcal{M}$ is a (closed) subspace, $\mathcal{Y}$ is a normed vector space and $T : \mathcal{X} \to \mathcal{Y}$ is continuous, then the mapping $\tilde{T}$ is bounded and $\|\tilde{T}\| = \|T\|$.*

Proof. Let $\pi : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ denote the quotient map and observe that $\tilde{T}\pi = T$. Since the quotient map π has norm 1 (see Problem 20.20), we see that $\|\tilde{T}\| \le \|T\|$. For the opposite inequality, let $0 < \epsilon < 1$ and choose $x \in \mathcal{X}$ such that $\|x\| = 1$ and $\|Tx\| > (1 - \epsilon)\|T\|$. Then $\|\pi(x)\| \le 1$ and

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

Letting ϵ go to zero finishes the proof. □

21.5. **Problems.**

Problem 21.1. 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 21.2. 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 $f \in \mathcal{Y}^*$,

$$|f(Tx)| \le C\|f\|\|x\|; \tag{8}$$

in which case $\|T\|$ is equal to the best possible C in (8).

Problem 21.3. 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 $\mathcal{M} + \mathbb{K}x$ is closed. Use this result to give another proof that every finite-dimensional subspace of $\mathcal{X}$ is closed.

Problem 21.4. Prove, if $\mathcal{M}$ is a *finite-dimensional* subspace of a Banach space $\mathcal{X}$, then there exists a closed subspace $\mathcal{N} \subset \mathcal{X}$ 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 21.1 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 21.5. Prove Proposition 21.8.

Problem 21.6. Prove every finite-dimensional Banach space is reflexive.

Problem 21.7. 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 $\Lambda : c_0 \to \ell^\infty$. As a corollary, conclude that c_0 is not reflexive. (Of course, saying $c_0 \neq \ell^\infty$ via the canonical embedding of Corollary 21.11 is much weaker than saying there is no isometric isomorphism between c_0 and ℓ^∞.)

Problem 21.8. 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 21.9. 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) 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 21.10. 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 $\mathcal{M} \subset \mathcal{X}$ 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 21.15.)

Problem 21.11. 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 21.12. 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 $\mathcal{M} \subset \ell^\infty$).

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 21.11.

Problem 21.13. Let $\mathcal{X}$ be a normed space and let $K \subset \mathcal{X}$ 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 $0 \le t \le 1$.) 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 $B = \mathit{ball}(\mathcal{X})$ 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 $\|x \pm y\| \le 1$.

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 $\mathit{ball}(\mathcal{X})$ and $\mathit{ball}(\mathcal{Y})$.

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 21.14. 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 21.13).

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 21.15. Let

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

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 21.13 and 21.14 .)

(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.)

22. The Baire Category Theorem and applications

Recall, a set D in a metric space X is *dense* if $\overline{D} = X$. Thus, D is dense if and only if D^c does not contain a nonempty open set, if and only if D has nontrivial intersection with every nonempty open set. A topological space X is called a *Baire space* if it has the following property: if $(U_n)_{n=1}^{\infty}$ is a countable sequence of open dense subsets of X, then the intersection $\cap_{n=1}^{\infty} U_n$ is dense in X.

Theorem 22.1 (The Baire Category Theorem)**.** *Every complete metric space X is a Baire space. In other words, if X is a complete metric space and if $(U_n)_{n=1}^{\infty}$ is a sequence of open dense subsets of X, then $\cap_{n=1}^{\infty} U_n$ is dense in X.*

Theorem 22.1 is true if X is a locally compact Hausdorff space and there are connections between the Baire Category Theorem and the axiom of choice.

A subset $E \subset X$ is *nowhere dense* if its closure has empty interior. Equivalently, $\overline{E}^c$ is open and dense. A set F in a metric space X is *first category* (or *meager*) if it can be expressed as the 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. For applications, the following corollary often suffices.

Corollary 22.2. *If X is a complete metric space, then X is not a countable union of nowhere dense sets; i.e., X is of second category in itself.*

Proof. Take complements and apply Theorem 22.1. ☐

Thus, the Baire property is used as a kind of pigeonhole principle: the "thick" Baire space X cannot be expressed as a countable union of the "thin" nowhere dense sets E_n. Equivalently, if X is Baire and $X = \bigcup_n E_n$, then at least one of the E_n is somewhere dense.

The following lemma should be familiar from advanced calculus.

Lemma 22.3. *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) $(\mathrm{diam}(C_n))$ converges to 0,

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

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

Moreover, if $x_n \in C_n$, then (x_n) converges to some x.

Proof of Theorem 22.1. 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. Fix such a W. Since 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 $W \cap U_1$. 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

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

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

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

The sequence of sets $(\overline{B(x_n, r_n)})$ satisfies the hypothesis of Lemma 22.3 and X is complete. Hence there is an $x \in X$ such that

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

□

We now give 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.)

Theorem 22.4 (The Principle of Uniform Boundedness (PUB))**.** *Suppose $\mathcal{X}, \mathcal{Y}$ are normed spaces and $\{T_\alpha : \alpha \in A\} \subset B(\mathcal{X}, \mathcal{Y})$ is a collection of bounded linear transformations from $\mathcal{X}$ to $\mathcal{Y}$. Let B denote the set*

$$B := \{x \in X : M(x) := \sup_\alpha \|T_\alpha x\| < \infty\}. \tag{9}$$

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 each integer $n \geq 1$ consider the set

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

Since each T_α is 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 $(\bar{E_N})^\circ$ 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 $x_0 - x \in E_N$ for all $\|x\| < r$. Thus, for every α and every $\|x\| < r$,

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

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 $\sup_\alpha \|T_\alpha\| \leq 2N/r < \infty$. □

Given a subset B of a vector space $\mathcal{X}$ and a scalar $s \in \mathbb{K}$, 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 $B \subset \mathcal{X}$ and $s \in \mathbb{K}$ 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 be open can be refined somewhat.

Lemma 22.5 (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 to 0;
(iii) there is an $s > 0$ such that $T(sB)$ contains an open ball centered to 0; and
(iv) $T(sB)$ contains an open ball centered to 0 for each $s > 0$.

Proof. This result is more or less immediate from the fact, for fixed z_0 and $r \in \mathbb{K}$, that the translation map $z \to z + z_0$ and the dilation map $z \to rz$ are continuous in a normed vector space. The implication (i) implies (ii) is immediate. The fact that $T(sB) = sT(B)$ for $s > 0$ readily shows (ii), (iii) and (iv) are equivalent.

To finish the proof it suffice to show (iv) implies (i). Accordingly, suppose (iv) holds and let $U \subset X$ 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 $B(x, s) \subset U$. The ball $sB = B(0, s) = B(x, r) - x$ is an open ball centered to 0. By hypothesis there is an $r > 0$ such that $B^{\mathcal{Y}}(0, r) \subset T(B(0, s))$. (Here we use $B^{\mathcal{Y}}$ to emphasize this ball is a subset of $\mathcal{Y}$.) By linearity of T,

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

Thus $T(U)$ is open. □

Theorem 22.6 (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. Observe that (i) follows immediately from (iii). To prove (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

$T(X) = \bigcup_{n=1}^{\infty} T(B(0,n))$. Since the range of T is assumed second category, there is an N such that $T(B(0,N))$ is second category and hence not nowhere dense. In other words, $\overline{T(B(0,N))}$ has nonempty interior. By scaling (see Lemma 22.5), $\overline{T(B(0,1))}$ has nonempty interior. Hence, there exists $p \in \mathcal{Y}$ and $r > 0$ such that $\overline{T(B(0,1))}$ contains the open ball $B^{\mathcal{Y}}(p,r)$. (Here the superscript $\mathcal{Y}$ is used to emphasize this ball is in $\mathcal{Y}$.) It follows that for all $\|y\| < r$,

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

In other words,

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

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

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

We will use the hypothesis that $\mathcal{X}$ is complete to prove $B^{\mathcal{Y}}(0, \frac{r}{4}) \subset T(B(0,1))$. 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 that $\sum_{j=1}^{\infty} y_j$ converges to y. Further, for each j there is an $x_j \in B(0, \frac{1}{2^j})$ such that $y_j = Tx_j$. Since

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

the series $\sum_{j=1}^{\infty} x_j$ converges to some $x \in B(0,1)$. It follows that $y = Tx$ by continuity of T. Consequently $y \in T(B(0,1))$ and the proof of (iii) is complete.

To prove (ii), let $\mathcal{M}$ denote the kernel of T and $\tilde{T}$ the mapping $\tilde{T} : \mathcal{X}/\mathcal{M} \to \mathcal{Y}$ determined by $\tilde{T}\pi = T$. By Lemma 21.17, $\tilde{T}$ is continuous and one-one. 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, $\tilde{T}$ is an open map. and consequently $\tilde{T}^{-1}$ 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, 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 22.7 (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).*

To state the final result of this section, we need a few more definitions. Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. The Cartesian product $\mathcal{X} \times \mathcal{Y}$ is a topological space in the *product topology*. A set is open in the product topology if and only if it can be written as a union of products of open sets. Alternately, a set O is open if and only if for each $z = (x, y) \in O$ there exists open sets $U \subset \mathcal{X}$ and $V \subset \mathcal{Y}$ such that $z \in U \times V \subset O$. It is not too hard to show that $\mathcal{X} \times \mathcal{Y}$ is metrizable (in fact the product topology can be realized by norming $\mathcal{X} \times \mathcal{Y}$, e.g. with the norm $\|(x, y)\| := \max(\|x\|, \|y\|)$). It is easy to see that a sequence $z_n = (x_n, y_n)$ converges in the product topology if and only if both (x_n) and (y_n) converge. Further, if $\mathcal{X}, \mathcal{Y}$ are both Banach spaces (complete), then $\mathcal{X} \times \mathcal{Y}$ is also complete and hence a Banach space. The space $\mathcal{X} \times \mathcal{Y}$ is equipped with the coordinate projections $\pi_{\mathcal{X}}(x, y) = x, \pi_{\mathcal{Y}}(x, y) = y$. It is clear from the definition of the product topology that these maps are continuous. (In fact the product topology is the coarsest topology such that the coordinate projections are continuous.)

Given a linear map $T : \mathcal{X} \to \mathcal{Y}$, its *graph* 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 22.2 gives an example where $G(T)$ is closed, but T is not continuous. On the other hand, if $\mathcal{X}, \mathcal{Y}$ are complete (Banach spaces), then $G(T)$ is closed if and only if T is continuous.

Theorem 22.8 (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. 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 22.7, $\pi_1^{-1} : X \to G(T)$ is bounded. Since π_2 is bounded, $\pi_2 \circ \pi_1^{-1} : X \to Y$ is continuous. To finish the proof, observe $\pi_2 \circ \pi_1^{-1}(x) = \pi_2(x, Tx) = Tx$.

∎

22.1. **Problems.**

Problem 22.1. Show that there exists a sequence of open, dense subsets $U_n \subset \mathbb{R}$ such that $m(\bigcap_{n=1}^{\infty} U_n) = 0$.

Problem 22.2. Consider the linear subspace $\mathcal{D} \subset c_0$ 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 22.3. Suppose $\mathcal{X}$ is a vector space equipped with two norms $\|\cdot\|_1, \|\cdot\|_2$ such that $\|\cdot\|_1 \leq \|\cdot\|_2$. Prove that if $\mathcal{X}$ is complete in both norms, then the two norms are equivalent.

Problem 22.4. 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 22.5. 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 22.6. Suppose that $\mathcal{X}$ is a vector space with a countably infinite basis. (That is, there is a linearly independent set $\{x_n\} \subset \mathcal{X}$ 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 $\mathcal{X}_n := \operatorname{span}\{x_1, \dots x_n\}$.)

Problem 22.7. 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 22.8. 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 22.9. 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 22.10. 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.

23. Hilbert spaces

23.1. **Inner products.** Let $\mathbb{K}$ denote either $\mathbb{C}$ or $\mathbb{R}$.

Definition 23.1. Let X be a vector space over $\mathbb{K}$. An *inner product* on X is a function $u : X \times X \to \mathbb{K}$ such that, for all $x, y, z \in X$ and all $\alpha, \beta \in \mathbb{K}$,

(i) $u(x, x) \geq 0$ and $u(x, x) = 0$ if and only if $x = 0$.
(ii) $u(x, y) = \overline{u(y, x)}$
(iii) $u(\alpha x + \beta y, z) = \alpha u(x, z) + \beta u(y, z)$.

Notice that items (ii) and (iii) together imply

(iv) $u(x, \alpha y + \beta z) = \overline{\alpha} u(x, y) + \overline{\beta} u(x, z)$.

$\triangleleft$

Remark 23.2. A function u satisfying only items (iii) and (23.1) is called a *bilinear form* (when $\mathbb{K} = \mathbb{R}$) or a *sesquilinear form* (when $\mathbb{K} = \mathbb{C}$). In this case, if (ii) is also satisfied then u is called *symmetric* ($\mathbb{R}$) or *Hermitian* ($\mathbb{C}$). A Hermitian or symmetric form satisfying $u(x, x) \geq 0$ for all x is called *positive semidefinite* or a *pre-inner product*.

Typically, u is written $\langle \cdot, \cdot \rangle$ so that $u(x, y) = \langle x, y \rangle$.

Finally, observe if u is a bilinear (resp. sesquilinear) form, then each $z \in X$ induces a linear functional on X defined by $x \mapsto u(x, z)$. $\diamond$

Theorem 23.3 (The Cauchy-Schwarz inequality)**.** *Suppose $\langle \cdot, \cdot \rangle$ is a pre-inner product on the vector space X.*

(i) For $x, y \in X$,

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle. \tag{10}$$

(ii) If $x, z \in X$ and $\langle z, z \rangle = 0$, then $\langle x, z \rangle = 0$.

(iii) The set

$$M = \{x \in X : \langle x, x \rangle = 0\}$$

is a subspace of X.

(iv) Let $Y = X/M$ and let $\pi : X \to Y$ denote the quotient map. The form $[p, q] = \langle x, y \rangle$ where $x, y \in X$ are any choices of vectors such that $\pi(x) = p$ and $\pi(y) = q$ is well defined and an inner product on Y.

Item (i) is known as the *Cauchy-Schwarz* inequality.

Proof. Fix $x, y \in X$. For $t \in \mathbb{R}$, let $\lambda = t\langle x, y\rangle$ and compute, using the nonnegativity assumption

$$\begin{aligned} 0 \leq& \langle x - \lambda y, x - \lambda y\rangle \\ =& \langle x, x\rangle - 2t|\langle x, y\rangle|^2 + t^2 |\langle x, y\rangle|^2 \langle y, y\rangle \\ =:& P(t). \end{aligned}$$

Since $P(t)$ is a nonnegative quadratic, its discriminant is nonpositive; i.e.,

$$|\langle x, y\rangle|^4 \leq \langle x, x\rangle \langle y, y\rangle |\langle x, y\rangle|^2$$

and the Cauchy-Schwarz inequality follows.

If $\langle z, z\rangle = 0$ and $x \in X$, then (i) immediately implies $\langle x, z\rangle = 0$, which proves (ii). In particular, if both $x, y \in M$ and $c \in \mathbb{C}$, then

$$\langle x + cy, x + cy\rangle = \langle x, x\rangle + c\langle y, x\rangle + \overline{c}\langle x, y\rangle + |c|^2 \langle y, y\rangle = 0$$

and so M is a subpace.

Item (iv) is an exercise in definition chasing. □

23.2. **Examples.**

$\mathbb{K}^n$: It is easy to check that the standard scalar product on $\mathbb{R}^n$ is an inner product; it is defined as usual by

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j \tag{11}$$

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

$$\langle z, w\rangle = \sum_{j=1}^{n} z_j \overline{w_j}. \tag{12}$$

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

$\ell^2(\mathbb{N})$: Let

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

We may view ℓ^2 as a subset of the vector space $\mathscr{S}$ of all sequences (with domain $\mathbb{N}$) with entrywise addition and scalar multiplication. Define, for sequences $a = (a_1, a_2, \ldots)$ and $b = (b_1, b_2, \ldots)$ in ℓ^2,

$$\sum_{n=1}^{\infty} a_n \overline{b_n}. \tag{13}$$

This series is seen to converge absolutely using the comparison test and the inequality $2|a_n b_n| \leq |a_n|^2 + |b_n|^2$. From here it is not hard to prove that ℓ^2 is closed under the vector space operations of $\mathscr{S}$ and is hence a vector space; and further that (13) defines an inner product, $\langle a, b\rangle = \sum a_n \overline{b_n}$, called the *standard inner product* on ℓ^2.

$L^2(\mu)$: Generalizing the previous example, let $(X, \mathscr{M}, \mu)$ be a measure space. Consider the set of all measurable functions $f : X \to \mathbb{K}$ such that

$$\int_X |f|^2 \, d\mu < \infty$$

The space $L^2(\mu)$ is defined to be this set, modulo the equivalence relation which declares f equivalent to g if $f = g$ almost everywhere. From the inequality $2|f\overline{g}| \le |f|^2 + |g|^2$ it follows that $L^2(\mu)$ is a vector space and that we can define the inner product on $L^2(\mu)$ by

$$\langle f, g\rangle = \int_X f\overline{g} \, d\mu. \tag{14}$$

23.3. **Norms.** Given a vector space X over $\mathbb{K}$ and a semi-inner product $\langle\cdot,\cdot\rangle$, define for each $x \in X$

$$\|x\| := \sqrt{\langle x, x\rangle}. \tag{15}$$

This quantity should act something like a "length" of the vector x. Clearly $\|x\| \ge 0$ for all x, and moreover we have:

Theorem 23.4. *Let X be a semi-inner product space over $\mathbb{K}$, with $\|\cdot\|$ defined by equation (15). Then for all $x, y \in X$ and $\alpha \in \mathbb{K}$,*

(a) $\|x + y\| \le \|x\| + \|y\|$
(b) $\|\alpha x\| = |\alpha|\|x\|$.

Thus $\|\cdot\|$ is a seminorm on X.

If $\langle\cdot,\cdot\rangle$ is an inner product, then also

(c) $\|x\| = 0$ *if and only if* $x = 0$,

and thus $\|\cdot\|$ is a norm on X.

Proof. For all $x, y \in X$ we have

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \|x\|^2 + 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \qquad (16) \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

where we have used the Cauchy-Schwarz inequality in (16). Taking square roots finishes the proof of item (a). Items (b) and (23.4) are left as exercises. □

When X is an inner product space, the quantity $\|x\|$ will be called the *norm* of x. Item (a) will be referred to as the *triangle inequality*. On $\mathbb{R}^n$,

$$\|x\| = (x_1^2 + \cdots + x_n^2)^{1/2},$$

is the usual Euclidean norm.

Lemma 23.5. *Let H be an inner product space equipped with the 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. □

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

Definition 23.6. 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 23.5 and the second from the positive definiteness of a norm.

Lemma 23.7. *If E is a subset of an inner product space H, then*

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

Theorem 23.8 (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. □

Theorem 23.9 (The Parallelogram Law)**.** *If H is an inner product space and $f, g \in H$, then*

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2). \tag{17}$$

Proof. From the definition of the norm coming from the inner product,

$$\|f \pm g\|^2 = \|f\|^2 + \|g\|^2 \pm 2\mathrm{Re}\, \langle f, g\rangle.$$

Now add. □

Subtracting, instead of adding, in the proof of the Parallelogram Law gives the polarization identity

$$\|f + g\|^2 - \|f - g\|^2 = 4\mathrm{Re}\, \langle f, g\rangle$$

in the case $\mathbb{K} = \mathbb{R}$.

Theorem 23.10 (The Polarization identity)**.** *If H is an inner product space over $\mathbb{R}$, then*

$$\langle f, g\rangle = \frac{1}{4}\left(\|f + g\|^2 - \|f - g\|^2\right). \tag{18}$$

If H is a complex inner product space, then

$$\langle f, g\rangle = \frac{1}{4}[\|f + g\|^2 - \|f - g\|^2 + \|f - ig\|^2 - \|f + ig\|^2] \tag{19}$$

Remark: An elementary (but slightly tricky) theorem of von Neumann says that if H is any vector space equipped with a norm $\|\cdot\|$ such that the parallelogram law (17) holds for all $f, g \in H$, then H is an inner product space with inner product given by formula (18) in the case of real scalars and formula (19) in the case of complex scalars. (The proof is simply to *define* the inner product by equation (18) or (19), and check that it is indeed an inner product.)

23.5. **Completeness.**

Definition 23.11. A *Hilbert space* over $\mathbb{K}$ is an inner product space X over $\mathbb{K}$ that is complete in the metric $d(x, y) = \|x - y\|$. (Here, as usual, $\mathbb{K}$ is either $\mathbb{C}$ or $\mathbb{R}$.) ◁

That the inner product spaces $\mathbb{K}^n$ are complete (and hence Hilbert spaces) is known from elementary analysis. (Note that the complex case follows from the real case, since the Euclidean norms are equal under the natural isomorphism $\mathbb{C}^n \cong \mathbb{R}^{2n}$.)

Theorem 23.12. *$L^2(\mu)$ is complete.*

Proof. We use Proposition 20.3. Accordingly suppose (f_k) is a sequence in $L^2(\mu)$ and $\sum_{k=1}^{\infty} \|f_k\| = B < \infty$. Define

$$G_n = \sum_{k=1}^{n} |f_k| \quad \text{and} \quad G = \sum_{k=1}^{\infty} |f_k|.$$

By the triangle inequality, $\|G_n\| \le \sum_{k=1}^n \|f_k\| \le B$ for all n. Thus, by the Monotone Convergence Theorem and the Cauchy-Schwarz inequality,

$$\int_X G^2\,d\mu = \lim_{n\to\infty}\int_X G_n^2\,d\mu \le B^2. \tag{20}$$

Thus G belongs to $L^2(\mu)$ and in particular $G(x) < \infty$ for almost every x. Hence, by the definition of G, the sum

$$\sum_{k=1}^{\infty} f_k(x)$$

converges absolutely for almost every x. Hence there is a measurable function f such that this sum converges a.e. to f. By construction, $|f| \le G$ and thus $f \in L^2(\mu)$. Moreover, for all n we have

$$\left|f - \sum_{k=1}^n f_k\right|^2 \le (2G)^2.$$

Equation (20) says that G^2 is integrable, so we can apply the Dominated Convergence Theorem to obtain

$$\lim_{n\to\infty}\left\|f - \sum_{k=1}^n f_k\right\|^2 = \lim_{n\to\infty}\int_X \left|f - \sum_{k=1}^n f_k\right|^2 d\mu = 0.$$

□

23.6. **Best approximation.** The results of the preceding subsection used only the inner product, but to go further we will need to invoke completeness. From now on, then, we work only with Hilbert spaces. We begin with a fundamental approximation theorem. Recall that if X is a vector space over $\mathbb{K}$, a subset $K \subseteq X$ is called *convex* if whenever $a, b \in K$ and $0 \le t \le 1$, we have $(1-t)a + tb \in K$ as well. (Geometrically, this means that when a, b lie in K, so does the line segment joining them.)

Theorem 23.13. *Suppose H is a Hilbert space. If $K \subseteq H$ 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. 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, $\|h-v\|^2 \ge d^2$. Hence, by the parallelogram law,

$$\begin{aligned}\left\|\frac{x-y}{2}\right\|^2 &= \frac{1}{2}\left(\|x-h\|^2 + \|y-h\|^2\right) - \left\|\frac{x+y}{2} - h\right\|^2 \\ &\le \frac{1}{2}\left(\|x-h\|^2 + \|y-h\|^2\right) - d^2.\end{aligned} \tag{21}$$

By assumption, 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 $n \ge N$, $\|k_n - h\|^2 < d^2 + \frac{1}{4}\epsilon^2$. By (21), if

$m, n \geq N$ then

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

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_0, and since K is closed, $k_0 \in K$. Since $(k_n - h)$ converges to $(k_0 - h)$ and $\|k_n - h\|$ converges to d it follows, by continuity of the norm, that $\|k_0 - h\| = d$.

It remains to show that k_0 is the unique element of K with this property. If $k' \in K$ and $\|k' - h\| = d$, then another application of $v = (k_0 + k')/2$ belongs to K, and $\|v - h\| \geq d$. By the parallelogram law again, equation (21) gives

$$0 \leq \left\| \frac{k_0 - k'}{2} \right\|^2 \leq \frac{1}{2}\left(\|k_0 - h\|^2 + \|k' - h\|^2\right) - d^2 = \frac{1}{2}(d^2 + d^2) - d^2 = 0.$$

Hence $k_0 = k'$. □

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. (Note that a subspace is always convex). What is significant is that in the case of a subspace, the minimizer k_0 has an elegant geometric description, namely, it is obtained by "dropping a perpendicular" from h to M. This is the content of the next theorem.

Since we will use the notation often, let us write $M \leq H$ to mean that M is a closed subspace of H.

Theorem 23.14. *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 $\lambda = t\overline{\langle h - f_0, f\rangle}$. Since $f_0 + \lambda f \in M$,

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

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, for all $f \in M$

$$\begin{aligned}
\|h - f\|^2 &= \|(h - f_0) + (f_0 - f)\|^2 && (22)\\
&= \|h - f_0\|^2 + \|f_0 - f\|^2 \quad \text{(why?)} && (23)\\
&\geq \|h - f_0\|^2. && (24)
\end{aligned}$$

Thus $\|h - f_0\| = \operatorname{dist}(h, M)$. □

Corollary 23.15. *If $M \leq H$, then $(M^\perp)^\perp = M$.*

Proof. By Lemma 23.7, $M \subset (M^\perp)^\perp$. Now suppose that $x \in (M^\perp)^\perp$. By Theorem 23.14 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 23.7, $x - m \in (M^\perp)^\perp$. Thus $x - m = 0$ by Lemma 23.7, 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 $E \subset \mathcal{N}$, then

$$\mathcal{M} = \cap_{\mathcal{N} \in \mathscr{E}} \mathcal{N}$$

is the *smallest closed subspace containing* E.

Corollary 23.16. *If E is a subset of H, then $(E^\perp)^\perp$ is equal to the smallest closed subspace of H containing E.*

Proof. The proof uses Lemma 23.7 freely. Evidently $E \subset (E^\perp)^\perp$. Further $(E^\perp)^\perp$ is a closed subspace. If M is a closed subspace containing E, then $E^\perp \supset M^\perp$ and hence $(E^\perp)^\perp \subset (M^\perp)^\perp = M$ by Corollary 23.15. □

Corollary 23.17. *A vector subspace E of a Hilbert space H is dense in H if and only if $E^\perp = (0)$.*

Lemma 23.18. *Suppose $M, N \leq H$. If M and N are orthogonal, then $M + N$ is closed.*

Given subspaces $M, N \leq H$ 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 $M \perp N$. Hence, $M \oplus N$ indicates that M, N are orthogonal closed subspaces 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 $m + n \in M + N$. □

Corollary 23.19. *If $M \leq H$, 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 23.14. Hence $x = m + (x - m) \in M \oplus M^\perp$. □

23.7. **The Riesz Representation Theorem.** 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,

$$|L_g(h)| = |\langle h, g\rangle| \leq \|g\|\|h\|.$$

So $\|L_g\| \leq \|g\|$, but in fact it is easy to see that $\|L_g\| = \|g\|$; just apply L_g to the unit vector $g/\|g\|$. 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 $\mathbb{K}^n$ takes the form L_g. More generally, every *bounded* linear functional on a Hilbert space has the form just described.

Theorem 23.20 (The Riesz RepresentationTheorem)**.** *If H is a Hilbert space and $\lambda : H \to \mathbb{K}$ 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 20.7, $\ker\lambda = \lambda^{-1}(\{0\})$ is a *proper, closed* subspace of H. Thus, by Theorem 23.14 (or Corollary 23.19) 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,

$$\begin{aligned} 0 &= \langle h - \lambda(h)f, f\rangle \\ &= \langle h, g\rangle - \lambda(h)\langle f, f\rangle. \end{aligned}$$

Thus $\lambda = L_g$, where $g = \frac{f}{\|f\|^2}$ and the proof is complete. □

23.8. **Duality for Hilbert space.** In the case $\mathbb{K} = \mathbb{R}$ the Riesz representation theorem identifies H^* with H. In the case $\mathbb{K} = \mathbb{C}$, 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. In particular, it is a simple matter to identify the adjoint of a bounded operator $T : H \to H$ as an operator $T^* : H \to H$. (See Theorem 21.14.) Namely, for $h \in H$, define T^*h as follows. Observe that the mapping $\lambda : H \to \mathbb{C}$ defined by $\lambda(f) = \langle Tf, h\rangle$ is (linear and) continuous. Hence, there is a vector T^*h such that

$$\langle Tf, h\rangle = \lambda(f) = \langle f, T^*h\rangle.$$

Conversely, if $S : H \to H$ is linear and

$$\langle Tf, h\rangle = \langle f, Sh\rangle$$

for all $f, h \in H$, then $S = T^*$. In particular T^* is completely determined by $\langle Tf, h\rangle = \langle f, T^*h\rangle$ for all $f, h \in H$. Further properties of the adjoints on Hilbert space appear in Problem 23.2.

A bounded operator T on a Hilbert space H is *self-adjoint* or *hermitian* if $T^* = T$. The proof of the following lemma uses the convenient fact that if T is abounded operator on H and if $\langle Th, g\rangle = 0$ for all $h, g \in H$, then $T = 0$.

Proposition 23.21. *Suppose T is a bounded self-adjoint operator on a Hilbert space H. If $\langle Th, h\rangle = 0$ for all $h \in H$, then $T = 0$.*

Proof. The proof involves a polarization argument. Given $h, g \in H$, we have

$$2\Re\langle Th, g\rangle = \langle T(h+g), h+g\rangle = 0.$$

Similarly,

$$2i\Im\langle Th, g\rangle = \langle T(h-ig), h-ig\rangle = 0.$$

Hence $\langle Th, g\rangle = 0$ and thus $T = 0$. □

Returning to Theorem 23.14, if $M \le H$ 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

$$Ph = f_0. \tag{25}$$

That is, Ph is characterized by $Ph \in M$ and $(h - Ph) \perp M$. If the space M needs to be emphasized we will write P_M for P.

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.

Theorem 23.22. *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) $Q = P_N$.

The mapping P is called the *orthogonal projection of H on M* and, for $h \in H$, the vector Ph is the *orthogonal projection* of h onto M.

Proof. In view of Corollary 23.19, $M + M^\perp = H$ and $M \cap M^\perp = (0)$ 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 $P^2 = P$). Note that $Ph = 0$ if and only if $h \in M^\perp$. Thus $\ker(P) = M^\perp$. Further, $P_{M^\perp} = I - P$. In particular, the ranges of P and $I - P$ are orthogonal.

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 $\|Ph\| \leq \|h\|$. In particular, P is a bounded operator on H. (See also Problem 22.10.)

Given $g, f \in H$, using orthogonality of the ranges of P and $I - P$,

$$\begin{aligned}\langle Pf, Pg\rangle =&\langle Pf, (I-P)g + Pg\rangle \\ =&\langle Pf, g\rangle\end{aligned}$$

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. It is also immediate that $I - P$ is the projection onto $M^\perp$.

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).

An easy computation shows that $Q(I - Q) = 0$. 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 $\|Qh\| \leq \|h\|$.

If (h_n) is a sequence from the range of Q which 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 $N^\perp = \ker(Q)$.

Finally, 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 $h \in H$, since $h - Qh = (I - Q)h \in \ker(Q) = N^\perp$, it follows that $Qh = P_N h$.

□

23.9. **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 23.23. *An orthonormal set is maximal if it is not contained in any larger orthonormal set. A maximal orthonormal set is called an (orthonormal) basis for H.* ◁

Proposition 23.24. *An orthonormal set E is maximal if and only if the only vector orthogonal to E is the zero vector. Equivalently, an orthonormal set E is maximal if and only if the span of E is dense in H.*

To prove the proposition use $H = \overline{E} \oplus E^{\perp}$.

Remark 23.25. 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 22.6. $\diamond$

Example 23.26. Here are some common examples of orthonormal bases.

(a) Of course the standard basis $\{e_1, \dots, e_n\}$ is an orthonormal basis of $\mathbb{K}^n$.

(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 $h : \mathbb{N} \to \mathbb{K}$ 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 functions $e_n(x) = e^{2i\pi nx}$. An easy exercise shows this set is orthonormal. Though not obvious, it is in fact a basis. (See Problem 23.7.)

$\triangle$

23.10. **Basis expansions.** Our ultimate goal in this section is to show that vectors in Hilbert space admit expansions as (possibly infinite) linear combinations of basis vectors.

Theorem 23.27. *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 $h \in H$,*

$$Ph = \sum_{j=1}^{n} \langle h, e_j \rangle e_j. \tag{26}$$

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 $1 \leq m \leq n$,

$$\begin{aligned} \langle h - g, e_m \rangle =& \langle h, e_m \rangle - \left\langle \sum_{j=1}^{n} \langle h, e_j \rangle e_j, e_m \right\rangle \\ &= \langle h, e_m \rangle - \sum_{j=1}^{n} \langle h, e_j \rangle \langle e_j, e_m \rangle \\ &= \langle h, e_m \rangle - \langle h, e_m \rangle = 0. \end{aligned}$$

It follows that $h - g$ is orthogonal to $\{e_1, \dots, e_n\}$ and hence to M. $\square$

Theorem 23.28 (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 for each n,* $\operatorname{span}\{f_1, \dots f_n\} = \operatorname{span}\{e_1, \dots e_n\}$.

Proof. Put $e_1 = f_1/\|f_1\|$. Assuming $e_1, \dots e_n$ have been constructed satisfying the conditions of the theorem, let $g_{n+1} = \sum_{j=1}^n \langle f_{n+1}, e_j\rangle e_j$, the orthogonal projection of f_{n+1} onto $M_n = \operatorname{span}\{e_1, \dots, e_n\}$. Thus g_{n+1} is orthogonal to M_n and not 0. Let $e_{n+1} = \frac{g_n}{\|g_{n+1}\|}$. □

It follows from the Gram-Schmidt process that H is finite dimensional as a Hilbert space if and only if H is finite dimensional as a vector space (and these dimensions agree).

Theorem 23.29 (Bessel's inequality)**.** *If $\{e_1, e_2, \dots\}$ is an orthonormal sequence in H, then for all $h \in H$*

$$\sum_{n=1}^\infty |\langle h, e_n\rangle|^2 \le \|h\|^2. \tag{27}$$

Proof. For $N \in \mathbb{N}^+$, let P_N denote the projection onto $M_N = \operatorname{span}(\{e_1, \dots, e_N\})$. Given $h \in H$, Theorem 23.27 and orthogonality gives,

$$\begin{aligned}\|h\|^2 &= \|P_N h + (I - P_N)h\|^2 \\ &= \|P_n h\|^2 + \|(I - P_N)h\|^2 \\ &\ge \|P_N h\|^2 \\ &= \sum_{j=1}^N |\langle h, e_j\rangle|^2.\end{aligned}$$

Since the inequality holds for all N, the proof is complete. □

Corollary 23.30. *If $E \subset H$ is an orthonormal set and $h \in H$, then $\langle h, e\rangle$ is nonzero for at most countably many $e \in E$.*

Proof. Fix $h \in H$ and a positive integer N, and define

$$E_N = \{e \in E : |\langle h, e\rangle| \ge \frac{1}{N}\}.$$

We claim that E_N is finite. If not, then it contains a countably infinite subset $\{e_1, e_2, \dots\}$. Applying Bessel's inequality to h and this subset, we get

$$\|h\|^2 \ge \sum_{n=1}^\infty |\langle h, e_n\rangle|^2 \ge \sum_{n=1}^\infty \frac{1}{N} = +\infty,$$

a contradiction. Hence,

$$\{e \in E : \langle h, e\rangle \ne 0\} = \bigcup_{N=1}^\infty E_N$$

is a countable union of finite sets, and therefore countable. □

Corollary 23.31. *Suppose $E \subset H$ is an orthonormal set and let $\mathcal{F}$ denote the collection of finite subsets of E. If $h \in H$, then*

$$\sup\{\sum_{e\in F} |\langle h, e\rangle|^2 : F \in \mathcal{F}\} \leq \|h\|^2. \tag{28}$$

At this point we pause to discuss convergence of infinite series in Hilbert space. 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.

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.) Of course absolute convergence implies unconditional convergence. For ordinary scalar series, or in a finite dimensional Hilbert space such as $\mathbb{K}^n$, unconditional convergence implies absolute convergence; however in infinite dimensional Hilbert space unconditional convergence need not imply absolute convergence as the example following Theorem 23.32 shows.

Theorem 23.32. *Suppose $E = \{e_1, e_2, \dots\} \subset H$ 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.

Further, if $h \in H$, then the series

$$\sum_{j=1}^{\infty} \langle h, e_j\rangle e_j \tag{29}$$

is unconditionally convergent and, letting g denote the (unconditional) sum,

$$\langle g, e_j\rangle = \langle h, e_j\rangle.$$

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 23.32 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.

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 if and only if for each $\epsilon > 0$ there is an N so that for all $m \geq n \geq N$,

$$\|s_m - s_n\|^2 = \sum_{j=n+1}^{m} |a_j|^2 < \epsilon \tag{30}$$

if and only if the series $\sum_{j=N+1}^{m} |a_j|^2 < \epsilon$ converges. Hence items (i) and (ii) are equivalent.

Now suppose $\sum_{j=1}^{\infty} |a_j|^2$ converges and let $\varphi : \mathbb{N} \to \mathbb{N}$. For numerical series, absolute convergence implies conditional convergence. Hence $\sum_{j=1}^{\infty} |a_{\varphi(j)}|^2$ converges and therefore, using the equivalence between item (ii) implies (i) it follows that the series

$$\sum_{k=1}^{\infty} a_{\varphi(k)}\, e_{\varphi(k)}$$

converges to some g', the limit of the partial sums

$$s'_n = \sum_{j=1}^{n} a_{\varphi(j)}\, e_{\varphi(j)}.$$

It remains to show $g' = g$.

Given $\epsilon > 0$, choose N so that (30) holds. Now choose $M \geq N$ so that

$$\{1, 2, \dots N\} \subseteq \{\varphi(1), \varphi(2), \dots \varphi(M)\}.$$

For $n \geq M$, let G_n be the symmetric difference of the sets $\{1, \dots n\}$ and $\{\varphi(1), \dots \varphi(n)\}$ (that is, their union minus their intersection). Since $n \geq M$, the set $G_n \subset \{N+1, N+2, \dots\}$. It follows that

$$\begin{aligned}
\|s_n - s'_n\|^2 &= \|\sum_{k \in G_n} \pm a_k e_k\|^2 && (31)\\
&= \sum_{k \in G_n} |a_k|^2 && (32)\\
&\leq \sum_{N+1}^{\infty} |a_j|^2 && (33)\\
&< \epsilon. && (34)
\end{aligned}$$

Thus $g' = g$. Hence item (ii) implies item (iii) and the proof of the first part of the theorem is complete.

For $h \in H$ Bessel's inequality 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 $g \in H$. To complete the proof, given $\epsilon > 0$, choose N so that if $n \geq N$, then

$$\|g - \sum_{j=1}^{n} \langle h, e_j \rangle e_j\| < \epsilon.$$

It follows that, by the Cauchy-Schwarz inequality, for $m \le n$,

$$|\langle g - \sum_{j=1}^{n} \langle h, e_j \rangle e_j, e_m \rangle|^2 \le \|g - \sum_{j=1}^{n} \langle h, e_j \rangle e_j\| \, \|e_m\| < \epsilon.$$

On the other hand,

$$\langle g - \sum_{j=1}^{n} \langle h, e_j \rangle e_j, e_m \rangle = \langle g, e_m \rangle - \langle h, e_m \rangle$$

and the desired conclusion follows. □

There is another notion of convergence in Hilbert space. Let I be an index set and let $\mathcal{F}$ denote the collection of finite subsets of I. Given $\{h_i : i \in I\}$, a collection of elements of H, the series

$$\sum_{i \in I} h_i$$

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 $F \subset G \in \mathcal{F}$,

$$\| \sum_{i \in G} h_i - h \| < \epsilon.$$

Proposition 23.33. *If E is an orthonormal subset of a Hilbert space H and $h \in H$, then*

$$\sum_{e \in E} \langle h, e \rangle e$$

converges (as a net). Moreover, if g is the limit (as a net), then, for each $e \in E$,

$$\langle g, e \rangle = \langle h, e \rangle.$$

If (h_j) is a sequence from H and

$$\sum_{j \in \mathbb{N}^+} h_j$$

converges (as a net) to some $h \in H$, then

$$\sum_{j=1}^{\infty} h_j$$

converges unconditionally to h.

Proof. Let $E_0 = \{e \in E : \langle h, e \rangle \neq 0\}$. From Corollary 23.30, E_0 is at most countable. Suppose E_0 is countable and choose an enumeration, $E_0 = \{e_1, e_2, \dots\}$. By Theorem 23.32, 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. Given $\epsilon > 0$, there is an N so that if $n \geq N$, then

$$|g - \sum_{j=1}^{n} \langle h, e_j\rangle e_j| < \epsilon$$

and, from Bessel's inequality,

$$\sum_{j=N}^{\infty} |\langle g, e_j\rangle|^2 < \epsilon^2.$$

Let $F = \{e_1, \ldots, e_N\}$. If $G \subset E$ is finite and $F \subset G$, then

$$\begin{aligned}\|g - \sum_{e\in G} \langle g, e\rangle\| \leq & \|g - \sum_{j=1}^{N} \langle g, e_j\rangle\| + \|\sum_{e\in G\setminus F} \langle g, e\rangle e\| \\ \leq & \epsilon + [\sum_{j=N}^{\infty} |\langle g, e_j\rangle|^2]^{\frac{1}{2}} \\ \leq & 2\epsilon.\end{aligned}$$

Hence $\sum_{e\in E} \langle h, e\rangle$ converges as a net to g. Further, by construction, $\langle g, e\rangle = \langle h, e\rangle$ for $e \in E_0$. On the other hand, if $e \notin E_0$, then, for each n,

$$\langle \sum_{j=1}^{n} \langle h, e_j\rangle e_j, e\rangle = 0,$$

and thus $\langle g, e\rangle = 0 = \langle h, e\rangle$.

The proof of the last part of the proposition is left as a (challenging) exercise. See Problem 23.14. □

Theorem 23.34. *If $E \subset H$ 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 and $h \in H$. In this case,

$$\sum_{e\in E} \langle h, e\rangle e$$

converges (as a net) to some $g \in H$ and moreover $\langle g, e\rangle = \langle h, e\rangle$ for all $e \in E$. Suppose $g \neq h$ and let $f = \frac{g-h}{\|g-h\|}$ so that f is a unit vector. If $e \in E$, then

$$\langle f, e\rangle = \frac{1}{\|g - h\|} \langle g - h, e\rangle = 0,$$

and thus E is not maximal. Hence (a) implies (b).

Now suppose (b) holds and let $h, g \in H$ be given. Given ϵ, choose a finite subset F of E such that if $F \subset G \subset E$, 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,

$$\begin{aligned} \epsilon > & \big| \langle h - \sum_{e \in G} \langle h, e \rangle e, g - \sum_{f \in G} \langle g, e \rangle e \big| \\ = & \big| \langle h, g \rangle - \sum_{e \in G} \langle h, e \rangle \, \langle e, g \rangle \big|. \end{aligned}$$

Hence (b) implies (c).

Item (d) follows from (c) by choosing $g = h$. Finally, suppose that (a) does not hold. In that case there exists a unit vector $h \in H$ such that h is orthogonal to E. In particular,

$$\sum_{e \in E} |\langle h, e \rangle| = 0$$

and (d) does not hold. □

Theorem 23.35. *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 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. This maximal element is then a basis (maximal orthonormal set). □

Remark 23.36. If H has a finite orthonormal basis $E = \{e_1, \dots, e_n\}$, then by Theorem 23.34(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 23.37. *Any two bases of a Hilbert space H have the same cardinality.*

In the proof we let $|S|$ denote the cardinality of the set S.

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 Corollary 23.30, 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

$$|F| = \left|\bigcup_{e\in E} F_e\right| \leq |E| \cdot \aleph_0 = |E|$$

where the last equality holds since E is infinite.

By symmetry, $|F| \leq |E|$ and the proof is complete. □

In light of this theorem, we make the following definition.

Definition 23.38. 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* or *separable Hilbert space.* ◁

23.11. **Weak convergence.** 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 23.39. 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 23.9). 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 23.40. *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 $\sum_j |\hat{h}_j|^2 \leq M$ (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,

$$\sum_{j=1}^{J} |\langle h_n, e_j \rangle|^2 \leq \|h_n\|^2 \leq M^2$$

by Bessel's inequality. Thus,

$$\sum_{j=1}^{J} |\hat{h}_j|^2 = \sum_{j=1}^{J} \lim_n |\langle h_n, e_j \rangle|^2 = \lim_n \sum_{j=1}^{J} |\langle h_n, e_j \rangle|^2 \leq M^2.$$

Thus $\sum_j |\hat{h}_j|^2 \leq M^2$ 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 23.32. By Theorem 23.34, $\|h\| \leq M$.

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,

$$|\langle h_n - h, g \rangle| \leq |\langle h_n - h, g_0 \rangle| + |\langle h_n - h, g_1 \rangle|.$$

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. $\square$

It turns out, if (h_n) converges to h weakly, then $\|h\| \leq \liminf \|h_n\|$ 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 23.9.

Theorem 23.41 (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.*

Theorem 23.41 holds without the separability hypothesis, but the proof is much simpler with the hypothesis.

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 23.11) □

23.12. **Problems.**

Problem 23.1. 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 23.2. (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 $a \in \mathbb{K}$, 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 23.3. 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 $E \subset H$ 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

$$S(a_0, a_1, a_2, \dots) = (0, a_0, a_1, \dots) \tag{35}$$

Prove S is an isometry, but not unitary. Compute S^* and SS^*.

Problem 23.4. For any set J, let $\ell^2(J)$ denote the set of all functions $f : J \to \mathbb{K}$ 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 23.3(c).)
b) Prove, if H is any Hilbert space, then H is isometrically isomorphic to $\ell^2(J)$ for some set J.

Problem 23.5. 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 23.6. (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 23.7. 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 $h_{mn}(x) := \mathbf{1}_{[m,m+1]}(x) g_n(x-m)$, $n \in \mathbb{N}, m \in \mathbb{Z}$ form an orthonormal basis for $L^2(\mathbb{R})$. (Thus $L^2(\mathbb{R})$ is separable.)

Problem 23.8. 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 23.9. (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 $\|h\| \leq \liminf \|h_n\|$.

Problem 23.10. Suppose H is countably infinite-dimensional (separable Hilbert space). Prove, if $h \in H$ and $\|h\| \leq 1$, 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 23.11. Prove Theorem 23.41.

Problem 23.12. 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 23.13. 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 23.14. 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.)