# Functional Analysis

Irena Swanson, Reed College, Spring 2011

This is work in progress. I am still adding, subtracting, modifying. Any comments, solutions, corrections are welcome.

**Table of contents:**

# 1 Overview, background

Throughout, $\mathbb{F}$ will stand for either $\mathbb{R}$ or $\mathbb{C}$. Thus $\mathbb{F}$ is a complete (normed) field.

[Conway's book assumes that all topological spaces are Hausdorff. Let's try to not impose that unless needed.]

Much of the material and inspiration came from Larry Brown's lectures on functional analysis at Purdue University in the 1990s, and some came from my Reed thesis 1987. Functional analysis is a wonderful blend of analysis and algebra, of finite-dimensional and infinite-dimensional, so it is interesting, versatile, useful.

I will cover Banach spaces first, Hilbert spaces second, as Banach spaces are more general.

# 2 Definition of Banach spaces

**Definition 2.1** *Let $X$ be a vector space over $\mathbb{F}$. A* **norm** *on $X$ is a function $\| \; \| : X \to [0, \infty)$ such that*

*(1) (Positive semidefiniteness) For all $x \in X$, $\|x\| = 0$ if and only if $x = 0$.*

*(2) $\|rx\| = |r| \, \|x\|$ for all $x \in X$ and all $r \in \mathbb{F}$.*

*(3) (Triangle inequality) For all $x, y \in X$, $\|x + y\| \leq \|x\| + \|y\|$.*

Is is easy to verify that for all $x, y \in X$, $\|x - y\| \leq \|x\| - \|y\|$, and that $| \, \|x\| - \|y\| \, | \leq \|x \pm y\|$.

Note that if $X$ is a normed vector space, then $X$ has a metric:

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

**Definition 2.2** *A vector space $X$ with a norm $\| \; \|$ is called a* **normed vector space** *or a* **normed linear space**. *In a normed linear space, a sequence $\{f_n\}$ is* **Cauchy (in the norm)** *if for all $\epsilon > 0$ there exists $N$ such that for all $m, n \geq N$, $\|f_n - f_m\| < \epsilon$. A sequence $\{f_n\}$* **converges (in the norm)** *to $f$ if for all $\epsilon > 0$ there exists $N$ such that for all $n \geq N$, $\|f_n - f\| < \epsilon$. A sequence $\{f_n\}$ in $X$* **is convergent (in the norm)** *if there exists $f$ in $X$ such that $\{f_n\}$ converges in the norm to $f$. A* **Banach space** *is a normed linear space in which every Cauchy sequence is convergent.*

Silly/important fact of the day: On February 2, 2011, MathSciNet lists 20331 publications with "Banach" in the title (and only 13539 publications with "Hilbert" in the title).

## 3 Examples of Banach spaces

Examples of Banach spaces are given in propositions in this section.

**Proposition 3.1** *For any* $p \in [1, \infty)$, $\mathbb{F}^n$ *is a Banach space under the* $\mathbb{L}^p$*-norm* $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$. *Why is this a norm? Hölder's inequality helps. For now I am assuming this, and I am not sure that I will get back to this. Presumably you have seen Hölder's inequality. And you probably have seen that the open ball topology determined by this norm is equivalent to the standard Euclidean topology (with* $p = 2$*), in which case, you should be able to prove easily that this is a Banach space. We will actually prove more general facts later.*

**Definition 3.2** *For any topological space* $X$*, let* $C(X) = \{f : X \to \mathbb{F}, f \text{ continuous}\}$.

Note that if X is compact, then the **uniform norm**, or the **sup norm**,

$$\|f\| = \sup\{|f(x)| : x \in X\}$$

is indeed a norm on $C(X)$. If $X$ is not compact, $\sup\{|f(x)| : x \in X\}$ may take on the value $\infty$, so we cannot have a norm.

**Proposition 3.3** *If* $X$ *is a compact topological space, then* $C(X)$ *is a Banach space.*

*Proof.* Let $\{f_n\}$ be a Cauchy sequence in $C(X)$. Then for all $\epsilon > 0$, there exists $N$ such that for all $m, n \geq N, \|f_m - f_n\| < \epsilon$. In particular, for all $x \in X, |f_m(x) - f_n(x)| \leq \|f_m - f_n\| < \epsilon$. It follows that $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{F}$. Since $\mathbb{F}$ is complete, we thus have that for all $x \in X$, there exists $f(x) \in \mathbb{F}$ such that $f_n(x) \to f(x)$. So $\{f_n\}$ converges pointwise to a function $f$.

We need to verify that $f \in C(X)$. It suffices to prove that for every open set $W$ in $\mathbb{F}$, its $f$-preimage $V$ in $X$ is an open subset of $X$. For that it suffices to prove that for every $x \in V$ there exists an open neighborhood $U$ of $x$ such that $U \subseteq V$. So let $x \in V$. Since $W$ is open in a metric space, there exists $\epsilon > 0$ such that $B(f(x), \epsilon) \subseteq W$. By the Cauchy sequence assumption there exists $N$ such that for all $m, n \geq N$, $\|f_m - f_n\| < \epsilon/3$. In particular, for all $y \in X$, $|f_m(y) - f_n(y)| \leq \|f_m - f_n\| < \epsilon/3$, and for any $m \geq N$,

$$|f_m(y) - f(y)| = |f_m(y) - \lim_{n\to\infty} f_n(y)| = \lim_{n\to\infty} |f_m(y) - f_n(y)| \leq \epsilon/3.$$

Since$f_N$ is continuous, $U = f_N^{-1}(B(f_N(x), \epsilon/3))$ is an open neighborhood of $x$ in $X$. For all $y \in U$,

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon,$$

so that $f(y) \in B(f(x), \epsilon) \subseteq W$. Thus $x \in U \subseteq V$, and $U$ is open, as desired.

Finally, we need to prove convergence in the norm. For all $m \geq N$ (with $N$ as above),

$$\begin{aligned} \|f_m - f\| &= \sup\{|f_m(x) - f(x)| : x \in X\} \\ &= \sup\{|f_m(x) - \lim f_n(x)| : x \in X\} \\ &\leq \epsilon. \end{aligned}$$

And we're done, without needing to use the Hausdorff condition. □

**Definition 3.4** *For any topological space $X$, define*

$$C_b(X) = \{f : X \to \mathbb{F} \text{ continuous and bounded}\}$$

*with the metric $\|f\| = \sup\{|f(x)| : x \in X\}$.*

**Proposition 3.5** *For any topological space $X$, $C_b(X)$ is a Banach space with the uniform norm. [If $X$ is compact, then $C_b(X)$ is a subset of the already established Banach space $C(X)$.]*

*Proof.* It is straightforward to see that $C_b(X)$ is a normed linear space. If $\{f_n\}$ is a Cauchy sequence in $C_b(X)$, then there exists $B$ such that for all $n$, $\|f_n\| \leq B$. Thus any pointwise limit of $\{f_n\}$ is also bounded. The rest of this proof is just like for $C(X)$. □

**Definition 3.6** *Let $X$ be a topological space. A function $f : X \to \mathbb{F}$ is said to have* **compact support** *if for all $\epsilon > 0$, the set $\{x : |f(x)| \geq \epsilon\}$ is compact. Define*

$$C_0(X) = \{f : X \to \mathbb{F} \text{ continuous with compact support}\}.$$

**Proposition 3.7** *For any topological space $X$, $C_0(X)$ is a closed linear subspace of $C_b(X)$, and hence a Banach space (under the uniform norm).*

*Proof.* We first show that $C_0(X) \subseteq C_b(X)$. Let $f \in C_0(x)$. For all $n$, define $S_n\{x \in X : |f(x)| \geq n\}$. By assumption, $S_n$ is compact, and since $f$ is continuous, $S_n$ is closed. Since $f$ is $\mathbb{F}$-valued, $\cap S_n = \emptyset$. Thus by HW 1.3, there exists $n$ such that $S_n = \emptyset$, which means that $f$ is bounded.

Clearly $C_0(X)$ is closed under scalar multiplication. Now let $f, g \in C_0(X)$. Let $\epsilon > 0$. Then

$$\{x : |f(x) + g(x)| \geq \epsilon\} \subseteq \{x : |f(x)| \geq \epsilon/2\} \cup \{x : |g(x)| \geq \epsilon/2\},$$

and since $f + g$ is continuous, this says that $\{x : |f(x) + g(x)| \geq \epsilon\}$ is a closed subset of a compact set, whence compact itself by Proposition 1, which proves that $C_0(X)$ is closed under addition. Since $0 \in C_0(X)$, it follows that $C_0(X)$ is a vector subspace of $C_b(X)$.

We next prove that $C_0(X)$ is a closed subset of $C_b(X)$. If $\{f_n\}$ is a Cauchy sequence in $C_0(X)$ in the norm, then $\{f_n\}$ is a Cauchy sequence in $C_b(X)$ in the norm. Since $C_b(X)$ is Banach, there exists $f \in C_b(X)$ such that $f_n \to f$ in the norm. Let $\epsilon > 0$. Then there exists $N$ such that for all $n \geq N$, $\|f_n - f\| < \epsilon/2$. If $|f(x)| \geq \epsilon$, then by the triangle inequality,

$$\epsilon \leq |f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq \|f - f_N\| + |f_N(x)| < \epsilon/2 + |f_N(x)|.$$

Thus $|f_N(x)| \geq \epsilon/2$. Thus $\{x : |f(x)| \geq \epsilon\} \subseteq \{x : |f_N(x)| \geq \epsilon/2\}$. Again by continuity of $f$, $\{x : |f(x)| \geq \epsilon\}$ is a closed subset of the compact set $\{x : |f_N(x)| \geq \epsilon/2\}$, so that by Proposition 1, $\{x : |f(x)| \geq \epsilon\}$ is compact. Thus $f \in C_0(X)$, and we have convergence in the norm: $\|f_n - f\| \to 0 \in C_0(X)$.

Finally, by Propositions 3.5 and 2, $C_0(X)$ is complete. □

**Proposition 3.8** *Here are some special cases of Banach spaces $C_b(X), C_0(X)$, all with the sup norm.*

*(1) Let $X$ be an arbitrary set with discrete topology. Then any function $X \to \mathbb{F}$ is continuous. We write*

$$\ell^\infty(X) = C_b(X) = \{\text{bounded functions } X \to \mathbb{F}\},$$
$$c_0(X) = C_0(X) = \{\text{functions } X \to \mathbb{F} \text{ with compact support}\},$$

*(2) If $X = \mathbb{N}$, we also write*

$$\ell^\infty = \ell^\infty(\mathbb{N}) = C_b(\mathbb{N}) = \text{the set of all bounded } \mathbb{F}\text{-valued sequences},$$
$$c_0(\mathbb{N}) = C_0(\mathbb{N}) = \text{the set of all sequences in } \mathbb{F} \text{ with limit } 0.$$

*(3) Let $X = \{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}\}$, with the topology inherited from $\mathbb{R}$. Then*

$$c = C_b(X) = C(X) = \text{the set of all convergent } \mathbb{F}\text{-valued sequences},$$
$$c_0 = C_0(X) = \text{the set of all convergent sequences in } \mathbb{F} \text{ with limit } 0.$$

**Definition 3.9** *For $a, b \in \mathbb{R}$, define*

$$C^k([a,b]) = \text{ the set of all functions } [a,b] \to \mathbb{F} \text{ with } k \text{ continuous derivatives.}$$

**Proposition 3.10** *$C^k([a,b])$ is a Banach space with the norm*

$$Norm\,(f) = \sum_{i=0}^{k} \left\| f^{(i)} \right\|,$$

*where the norm of the derivatives is the sup norm.*

*Proof.* It is straightforward to verify that $C^k([a,b])$ is a vector space and that $Norm$ is a norm.

Let $\{f_n\}$ be a Cauchy sequence in the norm $Norm$. Then for each $i = 0, \ldots, k$, $\{f_n^{(i)}\}$ is a Cauchy sequence in $C([a,b])$ in the sup norm. Thus by Proposition 3.3, for each $i = 0, \ldots, k$, there exists $g_i \in C([a,b])$ such that $\{f_n^{(i)}$ converges to $g_i$ uniformly on $[a,b]$. By Homework 1.2, $g_0, g_1, \ldots, g_{k-1}$ are differentiable, and $g_i' = g_{i+1}$. This proves that $\{f_n\}$ converges to $g_0$ in the norm $Norm$. This finishes the proof. □

**Remark 3.11** The Banach space above is an example of a Sobolev space. Here is another example: if $\Omega$ is an open subset of $\mathbb{R}^n$,

$$W^{p,m}(\Omega) = \{f \in \mathbb{L}^p(\Omega) : D^\alpha f \in \mathbb{L}^p(\Omega) \text{ for all } \alpha \in \mathbb{N}^n \text{ with } |\alpha| \le m\}$$

is a Sobolev space, with the norm

$$\|f\|_{p,m} = \left( \sum_{|\alpha| \le m} \|D^\alpha f\|_p^p \right)^{1/p} .$$

Sobolev spaces arise in the solutions sets of certain in partial differential equations. Perhaps somebody would want to do a project on Sobolev spaces?

## 4 $\mathbb{L}^p$ spaces

**Definition 4.1** *Let $X$ be a set. A collection $\Sigma$ of subsets of $X$ is called a* **$\sigma$-algebra** *if $\emptyset \in \Sigma$ and if $\Sigma$ is closed under complements and under countable unions. A function $\mu : X \to [0, \infty)$ is a* **measure** *if $\mu(\emptyset) = 0$ and if $\mu$ is countable additive, i.e., if for any pairwise disjoint $A_1, A_2, \ldots$ in $\Sigma$, $\mu(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$. A* **measure space** *is a triple $(X, \Sigma, \mu)$, where $\mu$ is a measure on the $\sigma$-algebra $\Sigma$ on a set $X$. A function $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is* **measurable** *if for all $\alpha \in \mathbb{R}$, $\{x \in X : f(x) > \alpha\} \in \Sigma$. A function $f :\to \mathbb{C}$ is* **measurable** *if $Re\, f$ and $Im f$ are both measurable.*

It is standard to show that the set of all measurable functions is closed under scalar multiplication. It is also sometimes (!) closed under addition. One problem to defining addition is that "infinity minus infinity" is not defined. But if $f, g : X \to \mathbb{R} \cup \{\pm\infty\}$ are measurable and $f + g$ is defined, then $f + g$ is measurable as well. This follows from the following for every real number $\alpha$:

$$\{x : f(x) + g(x) > \alpha\} = \bigcup_{r \in \mathbb{Q}} \left(\{x : f(x) > \alpha - r\} \cap \{x : g(x) > r\}\right),$$

which is a countable union of finite intersections of sets in $\Sigma$, so it is in $\Sigma$. This proves that $f + g$ is measurable.

It follows that for any measurable functions $f, g$ (to any codomain), as long as $f + g$ is defined, it is measurable.

**Definition 4.2** *For any measurable function $f : X \to \mathbb{R} \cup \{\pm\infty\}$, define $f_+, f_- : X \to [0, \infty]$ by*

$$f_+(x) = \max\{f(x), 0\}, \qquad f_-(x) = \max\{-f(x), 0\}.$$

It is easy to prove that $f = f_+ - f_-$, $|f| = f_+ + f_-$. The sum is well-defined. Certainly $f, |f|$ are measurable if each $f_+, f_-$ is, and we also have the implication that if $f$ is measurable, then $f_+, f_-, |f|$ are measurable. This follows from the following more general fact: for any measurable functions $f, g : X \to \mathbb{R} \cup \{\infty\}$, $\max\{f, g\}$ is measurable:

$$\{x : \max\{f, g\}(x) > \alpha\} = \{x : \max\{f(x), g(x)\} > \alpha\} = \{x : f(x) > \alpha\} \cup \{x : g(x) > \alpha\}.$$

It is straightforward to prove that if $f_n : X \to \mathbb{R} \cup \{\pm\infty\}$ are measurable, then $\sup\limits_n f_n$, $\inf_n f_n$ are measurable, (recall: $(\sup\limits_n f_n)(x) = \sup\{f_n(x) : n\}$) and hence also $\limsup_n f_n = \inf_n \sup\limits_{m \geq n} f_m$ and $\liminf_n f_n = \sup\limits_n \inf_{m \geq n} f_m$ are measurable. In particular, if $\{f_n\}$ converges pointwise, $\lim_n f_n$ is measurable.

**Definition 4.3** *Of special significance are* **simple** *functions: these are those functions that take on only finitely many values in $\mathbb{F}$.*

Note that if $f, g : X \to \mathbb{R} \cup \{\pm\infty\}$ are measurable, $\{x : f(x) > g(x)\} \in \Sigma$ because $\{x : f(x) > g(x)\} = \cup_{r \in \mathbb{Q}}(\{x : f(x) > r\} \cap \{x : g(x) \leq r\})$, which is a countable union of sets in $\Sigma$.

(What would be wrong with the following reasoning: $\{x : f(x) > g(x)\} \in \Sigma$ because $\{x : f(x) > g(x)\} = \{x : f(x) - g(x) > 0\}$.)

Every simple function can be written as $\sum_{i=1}^k c_i \chi_{E_i}$ for some $c_i \in \mathbb{F}$ and some sets $E_i \subseteq X$. For such a function to be measurable, we can write it so that all the $E_i$ are in $\Sigma$ and pairwise disjoint, and that all the $c_i$ are distinct. We can also impose that $\cup E_i = X$.

**Definition 4.4** *Two measurable functions $f, g$ with domain $X$ are $\mu$-***equivalent** *if $\{x \in X : f(x) \neq g(x)\}$ has $\mu$-value 0.*

It is straightforward to show that $\mu$-equivalence is an equivalence relation. If $f$ is $\mu$-equivalent to $g$, then for any scalar $c$, $cf$ is $\mu$-equivalent to $cg$, so that scalar multiplication is well-defined on $\mu$-equivalent classes. If $f$ and $g$ take on the values $\pm\infty$ on a set of measure 0, then $f + g$ is defined almost everywhere; and in addition of $f$ is $\mu$-equivalent to $f'$ and $g$ is $\mu$-equivalent to $g'$, then $f' + g'$ is defined almost everywhere and $f + g$ is $\mu$-equivalent to $f' + g'$.

**Definition 4.5** *The* **Lebesgue integral** *of a simple measurable function* $f = \sum_{i=1}^{k} c_i \chi_{E_i}$ *with all* $c_i \in [0,\infty)$ *is*

$$\int f d\mu = \sum_{i=1}^{k} c_i \mu(E_i).$$

*The* **Lebesgue integral** *of a measurable function* $f : X \to [0,\infty]$ *is*

$$\int f d\mu = \sup\{\int \varphi d\mu \mid \varphi : X \to [0,\infty) \text{ measurable and simple}, \varphi \le f\}.$$

*In a standard illogical terminology, we define the Lebesgue integral for many functions, but we say that a measurable* $f : X \to [0,\infty]$ *is* **Lebesgue-integrable** *if the integral is a real number. With that, a measurable* $f : X \to \mathbb{R} \cup \{\pm\infty\}$ *is* **Lebesgue-integrable** *if* $f_+$ *and* $f_-$ *are Lebesgue-integrable, and a measurable* $f : X \to \mathbb{C}$ *is* **Lebesgue-integrable** *if* $Ref$ *and* $Imf$ *are Lebesgue-integrable.*

One needs to verify that if $f$ is measurable and Lebesgue-integrable, then for any measurable $g$ that is $\mu$-equivalent to $f$, $g$ is Lebesgue-integrable, and $\int f d\mu = \int g d\mu$.

**Definition 4.6** *Let* $(X, \Sigma, \mu)$ *be a measure space. We will consider either the measurable functions with codomain* $\mathbb{R}$ *or* $\mathbb{R} \cup \{\pm\infty\}$, *in which case we will set* $\mathbb{F} = \mathbb{R}$, *or we will consider the measurable functions with codomain* $\mathbb{C}$, *in which case we will set* $\mathbb{F} = \mathbb{C}$. *For any real number* $p \ge 1$, *define* $\mathbb{L}^p = \mathbb{L}^p(X) = \mathbb{L}^p(X, \Sigma, \sigma) = \mathbb{L}^p_{\mathbb{F}}(X)$ *to be the set of all* $\mu$*-equivalent classes of measurable functions such that for each/any representative* $f$ *of the equivalence class,* $\int |f|^p d\mu$ *is finite. It is straightforward to verify that* $\mathbb{L}^p$ *is a vector space over* $\mathbb{F}$. *We define the norm:*

$$\|f\|_p = \left(\int |f|^p d\mu\right)^{1/p}.$$

*It is standard to write elements of* $\mathbb{L}^p$ *as functions, even though the elements are really equivalence classes of functions.*

We have essentially already verified that $\mathbb{L}^p$ is a vector space.

To prove that $\| \ \|_p$ is a norm, first of all, show that it is independent of the representative of the equivalence class, positive semi-definiteness is easy, so is the scalar property, but the triangle inequality requires a few more steps. Here is an outline:

(1) Since $p \ge 1$, there exists $q \ge 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$.

(2) If $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then for any $f \in \mathbb{L}^p$ and any $g \in \mathbb{L}^q$, $fg \in \mathbb{L}^1$ and $\|fg\|_1 \le \|f\|_p \|g\|_q$. (This is **Hölder's inequality**.) Outline of proof: The case $p = 1$ is trivial, so we may assume that $p > 1$. If $\|f\|_p = 0$, then without loss of generality $f = 0$, so that both sides are 0. So we may assume that $\|f\|_p \ne 0$, and similarly that $\|g\|_q \ne 0$. Define $\phi : (0,\infty) \to \mathbb{R}$ by $\phi(t) = \frac{1}{p}t - t^{1/p}$. Then $\varphi'(t) = \frac{1}{p}(1 - t^{1/p-1})$, $\varphi'(t) < 0$ for all $t \in (0,1)$ and $\varphi'(t) > 0$ for $t > 1$. By

calculus, $\varphi(t) \geq \varphi(1)$ for all $t$. In particular, $t^{1/p} \leq \frac{1}{p}t - \frac{1}{p} + 1 = \frac{1}{p}t + \frac{1}{q}$ for all $t$, and so for any non-negative $a$ and positive $b$, $(a/b)^{1/p} \leq \frac{1}{p}(a/b) + \frac{1}{q}$, which says that $a^{1/p}b^{1-\frac{1}{p}} \leq \frac{1}{p}a + \frac{1}{q}b$. Note that this latter inequality holds even if $b$ is zero, i.e., it holds for all non-negative real numbers $a$ and $b$. With $a = \frac{|f(x)|^p}{\|f\|_p^p}$, $b = \frac{|g(x)|^q}{\|g\|_q^q}$, this says that

$$\frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|_q^q}.$$

Now integrate both sides to get $\frac{\|fg\|_1}{\|f\|_p\|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$.

(3) (**Minkowski's inequality**.) Let $f, g \in \mathbb{L}^p$. Then $\|f+g\|_p \leq \|f\|_p + \|g\|_p$. This is clearly true of $p = 1$, so we may assume that $p > 1$. We may also assume that $\|f+g\|_p$ is non-zero. Set $q = \frac{1}{1-\frac{1}{p}} = \frac{p}{p-1} > 1$. Note that $\int (|f+g|^{p-1})^q d\mu = \int |f+g|^p d\mu < \infty$, so that $(f+g)^{p-1} \in \mathbb{L}^q$. Furthermore,

$$\left\|(f+g)^{p-1}\right\|_q = \left(\int (|f+g|^{p-1})^q d\mu\right)^{1/q} = \left(\int |f+g|^p d\mu\right)^{1/q} = \|f+g\|_p^{p/q} = \|f+g\|_p^{p-1}.$$

From $|f+g|^p \leq |f|\,|f+g|^{p-1} + |g|\,|f+g|^{p-1}$ and Hölder's inequality we get that

$$\|f+g\|_p^p \leq (\|f\|_p^p + \|g\|_p^p) \cdot \left\|(f+g)^{p-1}\right\|_q = (\|f\|_p^p + \|g\|_p^p) \cdot \|f+g\|_p^{p-1},$$

and after dividing through by $\|f+g\|_p^{p-1}$, we get Minkowski's inequality.

**Proposition 4.7** *$\mathbb{L}^p(X, \Sigma, \sigma)$ is a Banach space.*

*Proof.* We need to prove completeness in the norm $\|\ \|_p$. Let $\{f_n\}$ be a Cauchy sequence in $\mathbb{L}^p$. We need to find $f \in \mathbb{L}^p$ such that $f_n \to f$ in the norm. By taking a subsequence, without loss of generality for all $n$, $\|f_n - f_{n+1}\|_p < 2^{-n}$. Note that the sum $f_1 + \sum_{j=1}^{n-1}(f_{j+1} - f_j)$ (is defined a.e. and) equals $f_n$. Let $G_n = \sum_{j=1}^{n-1}|f_{j+1} - f_j|$. For all $n$, each $G_n$ is measurable, $0 \leq G_n \leq G_{n+1}$, and $\|G_n\|_p \leq \sum_{j=1}^{n-1}\|f_{j+1} - f_j\|_p \leq \sum_{j=1}^{n-1} 2^{-j} < 1$. By the Monotone Convergence Theorem,

$$\int (\lim_n G_n)^p d\mu = \int \lim_n G_n^p d\mu = \lim_n \int G_n^p d\mu \leq 1,$$

so that $\lim_n G_n \in \mathbb{L}^p$. In particular, $\sum_{j=1}^{\infty}|f_{j+1} - f_j| < \infty$ almost everywhere. Since absolute convergence implies convergence, the sequence $f_n = f_1 + \sum_{j=1}^{n-1}(f_{j+1} - f_j)$ converges almost everywhere, say to a function $f$. As $f$ is a limit of measurable functions, it is measurable. Since $|f| \leq |f_1| + \lim_n G_n$, we have that $\int |f|^p d\mu \leq \int (|f_1| + \lim_n G_n)^p d\mu = (\||f_1| + \lim_n G_n\|_p)^p \leq (\|f_1\| + 1)^p < \infty$, so that $f \in \mathbb{L}^p$. Furthermore, $f_n$ converges to $f$ in the $\mathbb{L}^p$ norm because $|f - f_n| \leq \lim_n G_n$, and since $\lim_n G_n$ is in $\mathbb{L}^p$, the Lebesgue Dominated Converegence Theorem applies:

$$(\lim_n \|f - f_n\|)^p = \lim_n \|f - f_n\|^p = \int \lim_n |f - f_n|^p d\mu = \int 0 d\mu = 0. \qquad \square$$

## 5 $\mathbb{L}^\infty$ spaces

Let $(X, \Sigma, \mu)$ be a measure space. Let $f$ be a measurable function $f$ (from $X$ to $\mathbb{C}$ or $\mathbb{R} \cup \{\pm\infty\}$, and $B \in \mathbb{R} \cup \{\infty\}$ such that for some set $E \in \Sigma$ of measure 0, $|f|\chi_{X \setminus E} \leq B$. Then clearly for any measurable $g$ that is $\mu$-equivalent to $f$, there exists $F \in \Sigma$ of measure 0, such that $|g|\chi_{X\setminus F} \leq B$. For such $f$ we define

$$\|f\|_\infty$$

to be the infimum of all such possible $B$. We just proved that $\|f\|_\infty = \|g\|_\infty$ whenever $f$ is $\mu$-equivalent to $g$.

**Definition 5.1** *Let $(X, \Sigma, \mu)$ be a measure space. Define $\mathbb{L}^\infty = \mathbb{L}^\infty(X) = \mathbb{L}^\infty(X, \Sigma, \sigma)$ to be the set of all $\mu$-equivalent classes of measurable functions for which $\|f\|_\infty < \infty$.*

**Theorem 5.2** *If $(X, \Sigma, \mu)$ be a measure space, then $\mathbb{L}^\infty(X)$ is a Banach space with the norm $\| \ \|_\infty$.*

*Proof.* Note that if $f \in \mathbb{L}^\infty$, then $f$ is bounded almost everywhere, and so by the discussion above Definition 4.5, $\mathbb{L}^\infty$ is closed under addition. The other properties of vector spaces are straightforward to establish for $\mathbb{L}^\infty$.

For the norm we only verify that triangle inequality: Let $B, C \in \mathbb{R}$ such that $f \leq B$ almost everywhere and $g \leq C$ almost everywhere. Then $f+g \leq B+C$ almost everywhere, so that $\|f+g\|_\infty \leq B + C$. Now we take the infimum over all such $B$ and $C$ to get that $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$.

That $\mathbb{L}^\infty$ is complete I leave for the exercises. □

## 6 Some spaces that are almost Banach, but aren't

**Definition 6.1** *Let $X = C^\infty([a,b])$ be the set of all functions $[a,b] \to \mathbb{R}$ that have derivatives of all orders (one sided derivatives at $a$ and $b$).*

*For each $k \in \mathbb{N}$, define $\|f\|_k = \sum_{i=1}^k \left\|f^{(i)}\right\|_\infty$. (Distinguish $\| \ \|_k$ from the $p$-norms in $\mathbb{F}^n$ or in $\mathbb{L}^p$.)*

By calculus, $X$ is a vector space. It is straightforward to verify that $\| \ \|_k$ is a norm for each $k \in \mathbb{N}$. (Question: why isn't $((f))_k = \left\|f^{(k)}\right\|_\infty$ a norm?)

The definition of convergence in $X$ **ought to be** as follows: $\{f_n\} \to f$ if and only if for all $i \geq 0$, $\{f_n^{(i)}\} \to f^{(i)}$ in the sup norm. However, no single norm $\| \ \|_k$ captures that. We can only say the quasi-norm thing: $\{f_n\} \to f$ if and only if for all $k \in \mathbb{N}$, $\{f_n\} \to f$ in the $\| \ \|_k$ norm.

**Definition 6.2** *Let $C(\mathbb{F}^n)$ be the set of all continuous functions $\mathbb{F}^n \to \mathbb{F}$.*

The definition of convergence in $C(\mathbb{F}^n)$ **ought to be** as follows: $\{f_n\} \to f$ if and only if for all compact subsets $K \subseteq \mathbb{F}^n$, $\{(f_n)|_K\} \to f|_K$ in the sup norm in $C(K)$. However, no single norm captures that. We can only say the quasi-seminorm thing: $((f))_K = \sup\{|f(x)| : x \in K\}$ is a seminorm on $C(\mathbb{F}^n)$, and $\{f_n\} \to f$ if and only if for all compact subsets $K \subseteq \mathbb{F}^n$, $\{(f_n)\} \to f|_K$ in the seminorm $((\ ))_K$.

**Remark 6.3** A normed vector space can have more than one norm on it, and it can be complete in one norm and not complete in another. We'll see an example in **'example-completenotcomplete'**.

## 7 Normed vector spaces are metric spaces

If $X$ is a normed vector space with metric $\|\ \|$, it is a metric space with the metric $d(x, y) = \|x - y\|$. This satisfies the property of metrics:

(1) $d(x, x) = \|x - x\| = \|0\| = 0$ for all $x$.

(2) If $d(x, y) = 0$, then $\|x - y\| = 0$, so that $x - y = 0$, so that $x = y$.

(3) For all $x, y, z$, $d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$.

Not every metric produces a norm, however! Namely, the **discrete metric** on a vector space (over a field with at least 3 elements) is given by $d(x, x) = 0$ and $d(x, y) = 1$ if $x \neq y$, and is a metric. However, the function $x \mapsto d(x, x)$ is not a norm as it does not obey the scalar rule.

In any case, a normed vector space $X$ has a norm on it, as well as the corresponding metric topology. Whereas a norm produces precise numbers, open sets are determined more loosely. This is reflected in Theorem 7.3 below. But first a lemma:

**Lemma 7.1** *Let $X$ be a vector space. Let $p$ be a norm and $q$ a seminorm on $X$. Suppose that for some positive real numbers $r, s$, $\{x : p(x) < r\} \subseteq \{x : q(x) < s\}$. Then for all $x \in X$,*

$$q(x) \leq \frac{s}{r} p(x).$$

*Proof.* If $q(x) = 0$, there is nothing to show. So we may assume that $q(x) \neq 0$. Then $x \neq 0$. As $p$ is a norm, then $p(x) \neq 0$. Let $\alpha \in (0, 1)$. Then $\alpha r \frac{x}{p(x)} \in \{y : p(y) < r\} \subseteq \{y : q(y) < s\}$, so that $q(\alpha r \frac{x}{p(x)}) < s$. In other words, $\alpha r q(x) < s p(x)$. Now we take the limit as $\alpha$ goes to 1 to obtain that $r q(x) \leq s p(x)$. This proves the lemma. ∎

**Lemma 7.2** *Let $\| \ \|$ and $(\!( \ )\!)$ be two norms on a vector space $X$. Then the open sets in the topology determined by $(\!( \ )\!)$ are open in the topology determined by $\| \ \|$ if and only if there exists a positive real number $B$ such that for all $x \in X$,*

$$B(\!(x)\!) \leq \|x\| .$$

*Proof.* Suppose that the open sets in the topology determined by $(\!( \ )\!)$ are open in the topology determined by $\| \ \|$. Then $\{x : (\!(x)\!) < 1\}$ is open in the topology determined by $\| \ \|$, so that there exists $r > 0$ such that $\{x : \|x\| < r\} \subseteq \{x : (\!(x)\!) < 1\}$. But then by Lemma 7.1 for all $x$, $(\!(x)\!) \leq \frac{1}{r} \|x\|$, so we may take $B = r$.

Now suppose that there exists a positive real number $B$ such that for all $x \in X$, $B(\!(x)\!) \leq \|x\|$. We need to prove that the $(\!( \ )\!)$-open sets are open in the $\| \ \|$ topology. So let $U$ be an $(\!( \ )\!)$-open set. Let $a \in U$. We need to prove that there exists a $\| \ \|$-open set containing $a$ that is contained in $U$. First of all, there exists $r > 0$ such that $\{x : (\!(x - a)\!) < r\} \subseteq U$. Now consider $V = \{x : \|x - a\| < Br\}$. This is open in the $\| \ \|$ topology and it contains $a$. If $x \in V$, then $B(\!(x - a)\!) \leq \|x - a\| < Br$, so that $(\!(x - a)\!) < r$, which implies that $x \in U$. Thus $V$ is an open set containing $a$ and contained in $U$. □

The following theorem is an immediate corollary:

**Theorem 7.3** *Let $\| \ \|$ and $(\!( \ )\!)$ be two norms on a vector space $X$. Then the topologies determined by these two norms are the same if and only if there exist positive real numbers $B$, $C$, such that for all $x \in X$,*

$$B(\!(x)\!) \leq \|x\| \leq C(\!(x)\!).$$ □

Topology also enables us to talk about continuous functions:

**Theorem 7.4** *Let $X$ and $Y$ be normed vector spaces. For a linear transformation $T : X \to Y$ the following are equivalent:*

*(1) $T$ is uniformly continuous.*
*(2) $T$ is continuous.*
*(3) $T$ is continuous at $0$.*
*(4) $T$ is continuous at some point.*
*(5) There exists $c > 0$ such that for all $x \in X$, $\|T(x)\| \leq c \|x\|$. (The two norms are one on $X$ and one on $Y$.)*

*Proof.* Trivially (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4), and (5) $\Rightarrow$ (3).

Assume (4). Thus $T$ is continuous at some point $a$. Let $\epsilon > 0$. Then there exists $\delta > 0$ such that for all $x$, $\|x - a\| < \delta$ implies that $\|T(x) - T(a)\| < \epsilon$. Now let $b \in X$ be arbitrary, and let $y \in X$ satisfy $\|y - b\| < \delta$. Then $\|(y - b + a) - a\| < \delta$, so that by continuity of $T$ at $a$, $\|T(y - b + a) - T(a)\| < \epsilon$. In other words, by linearity of $T$, $\|T(y) - T(b)\| < \epsilon$. This proves (1), so that (1) through (4) have been proved to be equivalent.

Now assume (3). Then there exists $\delta > 0$ such that for all $x$ with $\|x\| < \delta$, we have that $\|T(x)\| < 1$. In other words,

$$\{x \in X : \|x\| < \delta\} \subseteq \{x \in X : \|T(x)\| < 1\}.$$

Note that $x \mapsto \|T(x)\|$ is a seminorm. Thus by Lemma 7.1, for all $x \in X$, $\|T(x)\| \leq \frac{1}{\delta} \|x\|$. This proves (5), and finishes the proof of the theorem. □

**Remark 7.5** Not all linear functions are continuous. Let $X = \mathbb{L}^\infty(\mathbb{N}^+)$. Then $X$ contains $\{\frac{1}{n}\}_n$ and the sequences $e_n$ that are 1 in the $n$th entry and 0 elsewhere. The set $\{\{\frac{1}{n}\}_n, e_1, e_2, \ldots\}$ is linearly independent, so it can be extended into a vector space bases $B$. Define $T : X \to \mathbb{R}$ as $T(x)$ being the coefficient of $\{\frac{1}{n}\}_n$ in the writing of $x$ as a linear combination of elements of $B$. By uniqueness of such linear combinations, $T$ is linear. Set $x_m = \{\frac{1}{n}\}_n - e_1 - \frac{1}{2}e_2 - \cdots - \frac{1}{m}e_m = \{0, 0, \ldots, 0, \frac{1}{m+1}, \frac{1}{m+2}, \frac{1}{m+3}, \ldots\}$. Then $\|x_m\| = \frac{1}{m+1}$ and $\|T(x_m)\| = 1$, so that condition (5) in the theorem fails, and so $T$ is not continuous.

## 8 How to make new spaces out of existent ones (or not)

In this part we will look at various constructions one can do to ((complete) normed) vector spaces to create other ones (or to fail at it).

**Definition 8.1** *Let $I$ be an index set, and for each $i \in I$, let $B_i$ be a vector space over $\mathbb{F}$. The* **direct sum** *of the $B_i$ is*

$$\bigoplus_{i \in I} B_i = \{(b_i)_{i \in I} : b_i \in B_i \text{ and at most finitely many } b_i \text{ are non-zero}\}.$$

It is easy to verify that $\bigoplus_{i \in I} B_i$ is an $\mathbb{F}$-vector space with componentwise addition and scalar multiplication.

If each $B_i$ is normed with norm $\| \ \|_i$, we can make $\bigoplus_{i \in I} B_i$ normed with

$$\|(b_i)_i\| = \sum_{i \in I} \|b_i\|_i .$$

Even though $I$ may have huge cardinality, $\sum_{i \in I} \|b_i\|_i$ is a finite sum by the definition of elements of direct sums. It is easy to verify that $\| \ \|$ is a norm.

If $I$ is finite, and if each $B_i$ is a Banach space, then the direct sum is also a Banach space: any Cauchy sequence is a Cauchy sequence componentwise, so each component has a limit, and the limits of the components form an element of the direct sum that is the limit of the original Cauchy sequence.

Similar reasoning shows that if each $B_i$ is a Banach space and at most finitely many are non-zero vector spaces, then their direct sum is a Banach space.

In general, however, an infinite direct sum of non-trivial Banach spaces is not complete: Let $I_0$ be an infinite countable subset of $I$. We may assume that $I_0 = \mathbb{N}$. For each $i \in \mathbb{N}$, let $b_i \in B_i$ have norm $2^{-i}$. Let $c_n$ be the element of $\oplus_{i\in I} B_i$ that has $b_i$ in the $i$th component if $i \le n$, and has other components 0. Then it is easy to see that $\{c_n\}$ is a Cauchy sequence in $\oplus_{i\in I} B_i$, but that the limit does not exist (as it would have to have only finitely many non-zero entries).

COMMENT: If each $B_i$ has a vector space basis $S_i$, there is an obvious way of making elements of $S_i$ be thought of as elements of the direct sum. Then show that $\cup_{i\in I} S_i$ is a basis of $\oplus_{i\in I} B_i$.

**Definition 8.2** *Let $I$ be an index set, and for each $i \in I$, let $B_i$ be a vector space over $\mathbb{F}$. The* **direct product** *of the $B_i$ is*

$$\prod_{i\in I} B_i = \{(b_i)_{i\in I} : b_i \in B_i\}.$$

It is easy to verify that $\prod_{i\in I} B_i$ is an $\mathbb{F}$-vector space with componentwise addition and scalar multiplication.

If $I$ is finite, or if at most finitely many $B_i$ are non-zero vector spaces, then $\prod_{i\in I} B_i = \bigoplus_{i\in I} B_i$, so that has been handled above.

COMMENT: If each $B_i$ has a vector space basis $S_i$, there is an obvious way of making elements of $S_i$ be thought of as elements of the direct product. Show that if $I$ is infinite and if infinitely many $B_i$ are non-trivial vector spaces, then $\cup_{i\in I} S_i$ is NOT a basis of $\prod_{i\in I} B_i$.

Similarly, even if each $B_i$ is normed with norm $\| \ \|_i$, there is no norm on the direct product that would give convergence if and only if there is convergence in each component!

We thus now have two failed attempts and one successful try at making new ((complete) normed) vector spaces out of old ones.

Here is another partial success: For any $I$ and normed $B_i$ as above, and for any $p \ge 1$, consider

$$X = \{(b_i)_i \in \prod_{i\in I} B_i : \sum_{i\in I} \|b_i\|_i^p \text{ makes sense}\}.$$

What does "make sense" mean? Well, perhaps we have the following:

(1) $I$ is finite;
(2) or all but finitely many $b_i$ are zero;
(3) or all but countably many $b_i$ are zero, and the countable sum converges;
(4) or we need to talk about nets and convergence in nets (and we will not do that).

Here is a special case of $X$: $I = \mathbb{N}$, all $B_i$ are $\mathbb{F}$, and "makes sense" means that $\sum_{i\in I} \|b_i\|_i^p$ converges. But then $X = \ell^p$, with the familiar norm, and $X$ is even complete in the norm (by Proposition 4.7).

We have seen direct sums, direct products, and special subsets of direct products, some of which gave us Banach spaces and some of which didn't. Here is another iffy one (with a trivial proof):

**Proposition 8.3** *A linear subspace of a Banach space is Banach if and only if it is closed.*
□

Our experience with finite dimensional vector spaces perhaps makes us wonder whether non-closed linear subspaces can exist, but they do. Here is an example: Let $X = \ell^p$, and let $Y = X \cap \oplus_{i\in\mathbb{N}}\mathbb{F}$. As in the earlier discussion of direct sums, $Y$ is a vector space, but it is not complete, and it is not closed in $X$.

For the especially good results on finite-dimensional vector spaces, see Section 9.

**Definition 8.4** *Let $X$ and $Y$ be normed vector spaces. Define*

$$B(X,Y) = \text{ the set of all continuous linear transformations } X \to Y.$$

It is clear that $B(X,Y)$ is a vector space. It is a normed space with the following:

$$\|T\| = \inf\{c > 0 : \text{ for all } x \in X,\ \|T(x)\| \le c\,\|x\|\}.$$

Note that with this definition, for all $x \in X$,

$$\|T(x)\| \le \|T\|\,\|x\|\,.$$

**Proposition 8.5** *The function $\|\ \|$ is a norm on $B(X,Y)$.*

*Proof.* For all $n \in \mathbb{N}$, $\|0(x)\| = 0 \le \frac{1}{n}\|x\|$, so that $\inf\{c > 0 : \text{ for all } x \in X,\ \|T(x)\| \le c\,\|x\|\} \le \inf\{\frac{1}{n} > 0 : n \in \mathbb{N}\} = 0$. Thus the norm of the zero linear transformation is 0.

Suppose that for some $T$, $\inf\{c > 0 : \text{ for all } x \in X,\ \|T(x)\| \le c\,\|x\|\} = 0$. Then for all $n \in \mathbb{N}$ and for all $x \in X$, $\|T(x)\| \le \frac{1}{n}\|x\|$. Thus if we fix $x$ and let $n$ go to infinity, we get that $\|T(x)\| = 0$ for each $x$, and since this $\|\ \|$ is a norm on $Y$, we have that $T(x) = 0$ for all $x \in X$, so that $T$ is the zero linear transformation.

It is straightforward to prove that for all $\alpha \in \mathbb{F}$ and all $T$, $\|\alpha T\| = |\alpha|\,\|T\|$.

Now let $T, S \in B(X,Y)$. Then for all $x \in X$,

$$\|(T+S)(x)\| = \|T(x) + S(x)\| \le \|T(x)\| + \|S(x)\| \le \|T\|\,\|x\| + \|S\|\,\|x\| = (\|T\| + \|S\|)\,\|x\|\,,$$

whence $\|T+S\| = \inf\{c > 0 : \text{ for all } x \in X,\ \|(T+S)(x)\| \le c\,\|x\|\} \le \|T\| + \|S\|$. This proves the triangle inequality. □

It is convenient to have alternate formulations of this norm:

**Proposition 8.6** *Let $X$ and $Y$ be normed vector spaces. Let $T \in B(X,Y)$. Then the following numbers are the same:*

*(1)* $n_1 = \inf\{c > 0 :$ *for all* $x \in X$, $\|T(x)\| \le c\|x\|\}$.
*(2)* $n_1' = \inf\{c > 0 :$ *for all* $x \in X \setminus \{0\}$, $\|T(x)\| \le c\|x\|\}$.
*(3)* $n_2 = \inf\{c > 0 :$ *for all* $x \in X$ *with* $\|x\| \le 1$, $\|T(x)\| \le c\}$.
*(4)* $n_3 = \inf\{c > 0 :$ *for all* $x \in X$ *with* $\|x\| < 1$, $\|T(x)\| \le c\}$.
*(5)* $n_4 = \sup\{\|T(x)\| : x \in X, \|x\| \le 1\}$.
*(6)* $n_4' = \sup\{\|T(x)\| : x \in X \setminus \{0\}, \|x\| \le 1\}$.
*(7)* $n_5 = \sup\{\|T(x)\| : x \in X, \|x\| < 1\}$.
*(8)* $n_5' = \sup\{\|T(x)\| : x \in X \setminus \{0\}, \|x\| < 1\}$.

*Proof.* I leave the proof of $n_1 = n_1'$, $n_4 = n_4'$, $n_5 = n_5'$ and $n_4 = n_5$ to you.

If $x \in X$ has $\|x\| \le 1$ (or $< 1$), then $\|T(x)\| \le n_1\|x\| \le n_1$, so that $n_2, n_3 \le n_1$. If $x \in X$ is non-zero, then $\|T(x)\| = \|x\|\,\|T(x/\|x\|)\| \le \|x\| n_2$, which proves that $n_1' \le n_2$. If $\alpha \in (0,1)$, then in addition $\|T(x)\| = \frac{\|x\|}{\alpha}\|T(\alpha x/\|x\|)\| \le \frac{\|x\|}{\alpha} n_3$, so that in the limit as $\alpha \to 1$, $\|T(x)\| \le \|x\| n_3$, which proves that $n_1' \le n_3$.

Let $x \in X$ satisfy $\|x\| \le 1$. Then $\|T(x)\| \le n_2$, which proves that $n_4 \le n_2$. Also, $\|T(x)\| \le n_4$, which proves that $n_2 \le n_4$. □

**Theorem 8.7** *If $X$ is a normed vector space and $Y$ is a Banach space, then $B(X,Y)$ is a Banach space.*

*Proof.* We have seen that $B(X,Y)$ is a normed vector space. Now let $\{T_n\}$ be a Cauchy sequence in $B(X,Y)$. Then for all $x \in X$ and all $n, m \in \mathbb{N}$,

$$\|T_n(x) - T_m(x)\| \le \|T_n - T_m\|\,\|x\|,$$

so that $\{T_n(x)\}$ is a Cauchy sequence in $Y$. Since $Y$ is complete, there exists $T(x) \in Y$ (depending on $x$) such that $\{T_n(x)\}$ converges to $T(x)$ in the norm.

The function $T : X \to Y$ is linear as for all $a, b \in \mathbb{F}$ and all $x, y \in X$,

$$T(ax+by) = \lim_n T_n(ax+by) = \lim_n (aT_n(x)+bT_n(y)) = a\lim_n T_n(x) + b\lim_n T_n(y) = aT(x)+bT(y).$$

$T$ is continuous: Since $\{T_n(x)\}$ is a Cauchy sequence, it is bounded. So there exists $c > 0$ such that for all $n$, $\|T_n\| \le c$. Then for all $x \in X$, $\|T(x)\| = \|\lim_n T_n(x)\| = \lim_n \|T_n(x)\|$ by the definition of limits in the norm in $Y$, whence

$$\|T(x)\| \le \lim_n \|T_n\|\,\|x\| \le \lim_n c\,\|x\| = c\,\|x\|,$$

so that by Theorem 7.4, $T$ is continuous.

$\{T_n\}$ converges to $T$ (in the norm on $B(X,Y)$): Let $\epsilon > 0$. Since $\{T_n\}$ is a Cauchy sequence, there exists $N$ such that for all $m, n \ge N$, $\|T_n - T_m\| < \epsilon/2$. Let $x \in X$ be non-zero. Since $\{T_n(x)\}$ is a Cauchy sequence, there exists $N'(x) \geq N$ such that for all $n \geq N'(x)$, $\|T_n(x) - T(x)\| < \epsilon \|x\|/2$. Then for all $n \geq N$, choose $m \geq N'(x)$, and then

$$\begin{aligned}\|T_n(x) - T(x)\| &\leq \|T_n(x) - T_m(x)\| + \|T_m(x) - T(x)\| \\ &< \|T_n - T_m\| \|x\| + \epsilon \|x\|/2 \\ &\leq \epsilon \|x\|.\end{aligned}$$

Thus for all $n \geq N$ and all (zero and non-zero) $x$, $\|T_n(x) - T(x)\| \leq \epsilon \|x\|$, so that $\|T_n - T\| \leq \epsilon$. □

**Corollary 8.8** *If $X$ is a normed vector space, then $B(X, \mathbb{F})$ is a Banach space.* □

**Definition 8.9** *For any normed vector space $X$ over the field $\mathbb{F}$,* **the dual space of X** *is $B(X, \mathbb{F})$, and it is denoted $X^*$.*

We proved that the dual space of a normed vector space is a Banach space.

We now move to another construction of ((complete) normed) vector spaces from existing ones.

**Definition 8.10** *Let $X$ be a vector space and $M$ a linear subspace. For any $x \in X$, define*

$$x + M = \{x + m : m \in M\}.$$

It is easy to see the following:

**Lemma 8.11** *$x + M = y + M$ if and only if $x - y \in M$.* □

**Definition 8.12** *Let $M$ be a linear subspace of an $\mathbb{F}$-vector space $X$. The* **quotient (space)** *of $X$ by $M$ is the set*

$$X/M = \{x + M : x \in X\}.$$

*On this set we define $+ : (X/M) \times (X/M) \to (X/M)$ and $\cdot : \mathbb{F} \times (X/M) \to (X/M)$ as*

$$\begin{aligned}(x + M) + (y + M) &= (x + y) + M, \\ r \cdot (x + M) &= (rx) + M.\end{aligned}$$

We first need to establish that $+$ and $\cdot$ are well-defined. Namely, we need to establish that $(x + y) + M = (x' + y') + M$ and that $(rx) + M = (rx') + M$. But the assumptions mean that $x - x', y - y' \in M$, so that $(x + y) - (x' + y'), (rx) - (rx') \in M$ as $M$ is an $\mathbb{F}$-vector space, whence the conclusions hold, so that $+$ and $\cdot$ are well-defined.

With that, it is straightforward to establish that $X/M$ is a vector space over $\mathbb{F}$. The zero vector is of course $0 + M = M$.

**Definition 8.13** *If $X$ is a normed vector space and if $M$ is a closed vector subspace, we define the norm on $X/M$ by*

$$\begin{aligned}\|x+M\| &= \inf\{\|x+m\| : m \in M\}\\ &= \inf\{\|y\| : y \in x+M\}\\ &= \inf\{\|x-m\| : m \in M\}\\ &= \textit{distance}(x, M).\end{aligned}$$

This is indeed a norm:

(1) It is real-valued, taking on only non-negative values.
(2) $\|0+M\| \leq \|0+0\| = 0$, so $\|0+M\| = 0$.
(3) If $\|x+M\| = 0$, then for all $n \in \mathbb{N}$, there exists $m_n \in M$ such that $\|x+m_n\| \leq 1/n$. Thus 0 is a limit point of $x+M$, so that $-x$ is a limit point of $(-x)+x+M = M$. Since $M$ is closed, $-x \in M$, whence $x \in M$, and $x+M = 0+M$ is the zero vector in $X/M$.
(4) Certainly $\|c(x+M)\| = |c|\,\|x+M\|$ for all $c \in \mathbb{F}$ and all $x \in X$.
(5) Let $x_1, x_2 \in X$. Let $\epsilon > 0$. Choose $m_1, m_2 \in M$ such that $\|x_1+m_1\| < \|x_1+M\| + \epsilon/2$, $\|x_2+m_2\| < \|x_2+M\| + \epsilon/2$. Then

$$\begin{aligned}\|(x+M)+(y+M)\| &= \|(x+y)+M\|\\ &\leq \|x+y+m_1+m_2\|\\ &\leq \|x+m_1\| + \|y+m_2\|\\ &< \|x_1+M\| + \|x_2+M\| + \epsilon.\end{aligned}$$

As we let $\epsilon$ go to 0, this shows that that $\|(x+M)+(y+M)\| \leq \|x_1+M\| + \|x_2+M\|$.

Thus we do have a norm on $X/M$.

**Theorem 8.14** *Let $X$ be a Banach space and $M$ a closed linear subspace. Then $X/M$ is a Banach space.*

*Proof.* We use Homework 4, Problem 1. Let $\{x_n+M\}$ be a sequence in $X/M$ such that $\sum_n \|x_n+M\|$ converges.

Recall: $\|x+M\| = \inf\{\|y\| : y \in x+M\}$. Thus for all $n$, let $y_n \in x_n+M$ such that $\|x_n+M\| \leq \|y_n\| < \|x_n+M\| + 2^{-n}$. Thus $\sum_n \|y_n\|$ converges in $\mathbb{F}$. By Homework 4, Problem 1, since $X$ is complete, there exists $y \in X$ such that $\lim_n \sum_{i=1}^n y_i = y$ (in the norm). Then for all $n$,

$$\left\|\sum_{i=1}^n (x_i+M) - (y+M)\right\| = \left\|\sum_{i=1}^n (y_i+M) - (y+M)\right\| = \left\|(\sum_{i=1}^n y_i - y) + M\right\| \leq \left\|\sum_{i=1}^n y_i - y\right\|,$$

which goes to 0 as $n$ goes to infinity. □

We combine finite direct sums and quotients: if $X$ and $Y$ are Banach spaces, then $X \oplus 0$ is a closed subspace of $X \oplus Y$, and $(X \oplus Y)/(X \oplus 0)$ is naturally isomorphic to $Y$.

## 9 Finite-dimensional vector spaces are special

**Theorem 9.1** *If $X$ is a finite-dimensional vector space over $\mathbb{F}$, then any two norms define equivalent topologies.*

*Proof.* Without loss of generality $X \neq 0$. Let $\| \ \|$ be a norm on $X$. Let $\{e_1, \ldots, e_n\}$ be a vector space basis of $X$. Then for all $x \in X$, there are unique $x_i \in \mathbb{F}$ such that $x = \sum_i x_i e_i$. It is easy to prove that $\|x\|_\infty = \sum_i |x_i|$ is a norm on $X$. It suffices to prove that the norms $\| \ \|_\infty$ and $\| \ \|$ produce equivalent topologies. Let $C = \max\{\|e_i\| : i\}$. By the definition of norms,

$$\|x\| = \left\| \sum_i x_i e_i \right\| \leq \sum_i |x_i| \, \|e_i\| \leq C \sum_i |x_i| = C \, \|x\|_\infty .$$

Thus if $U$ is open in the topology determined by $\| \ \|$, by Lemma 7.2 it is open in the topology determined by $\| \ \|_\infty$.

Consider the set $B = \{x : \|x\|_\infty \leq 1\}$. This is a closed and bounded subset of a finite-dimensional metric space, hence is compact (by Proposition 2). Let $S$ be an open cover of $B$ in the topology determined by $\| \ \|$. By what we have already proved, $\mathcal{S}$ is an open cover of $B$ in the topology determined by $\| \ \|_\infty$, so that $\mathcal{S}$ has a finite subcover of $B$. This proves that $B$ is compact in both topologies. Similarly, $\{x : \|x\|_\infty = 1\}$ is compact in both topologies. As the $\| \ \|$-topology is metric, it is Hausdorff, so that it is closed. which means that $\{x : \|x\|_\infty < 1\}$ is open in the $\| \ \|$-topology as a subset of $B$, which means that there exists an open set $U$ in the $\| \ \|$-topology such that $U \cap B = \{x : \|x\|_\infty < 1\}$. In particular, there exists $r > 0$ such that $\{x : \|x\| < r\} \subseteq U$. This implies that for all $x$,

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

Of course, we want the implication $\|x\| < r \implies \|x\|_\infty < 1$. Without loss of generality $x \neq 0$. Let $\|x\| < r$. If $\|x\|_\infty \leq 1$, then by the displayed implication, $\|x\|_\infty < 1$. So we may assume that $\|x\|_\infty > 1$. Then $x / \|x\|_\infty \in B$ and $\|x / \|x\|_\infty\| = \|x\| / \|x\|_\infty < r / \|x\|_\infty < r$, so that by the displayed implication, $1 = \|x / \|x\|_\infty\|_\infty < 1$, which is a contradiction.

Thus we are done by Theorem 7.3. □

**Corollary 9.2** *A finite-dimensional vector subspace of a normed vector space is closed.*

*Proof.* Type ... □

**Theorem 9.3** *Any linear transformation from a finite-dimensional normed vector space to a normed vector space is continuous.*

*Proof.* Let $X, Y$ be normed vector spaces with $X$ finite-dimensional. Let $T : X \to Y$ be linear. Let $\{e_1, \ldots, e_n\}$ be a vector space basis of $X$. Let $C = \max\{\|T(e_i)\| : i = 1, \ldots, n\} + 1$. Then $C$ is a positive real number, and for all $x = \sum_i x_i e_i \in X$,

$$\|T(x)\| = \left\| \sum_i x_i T(e_i) \right\| \le \sum_i |x_i| \, \|T(e_i)\| \le \sum_i |x_i| C = C \, \|x\|_\infty .$$

Thus by Theorem 7.4, $T$ is continuous. □

**Corollary 9.4** *Any linear transformation from a finite-dimensional normed vector space to $\mathbb{F}$ is continuous.* □

... and from linear algebra we then know that $(\mathbb{F}^n)^* = B(\mathbb{F}^n, \mathbb{F}) \cong \mathbb{F}^n$.

**Theorem 9.5** *Let $X$ be a complete normed vector space, $M$ a closed subspace and $N$ a finite-dimensional subspace. Then $M + N = \{m + n : m \in M, n \in N\}$ is a closed subspace of $X$.*

*Proof.* (Part of the proof means deciphering the notation.) The image $(N + M)/M$ of $N$ in $X/M$ is a finite-dimensional vector subspace of $X/M$, and since $X/M$ is normed, by the corollary above we have that $(N + M)/M$ is a closed subspace of $X/M$.

Now let $\{n_i + m_i\}_i$ be a sequence in $N + M$ that converges to $x \in X$. We need to prove that $x \in M + N$. Clearly $\{n_i + m_i + M\}_i$ is a sequence in $(N+M)/M$ that converges to $x + M \in X/M$. Since $(N + M)/M$ is closed in $X/M$, necessarily $x + M = n + M$ for some $n \in N$, or in other words, $x = n + m$ for some $m \in M$. □

Well, why do we need to assume that $N$ is finite-dimensional? Isn't the sum of two closed linear subspaces always closed? (The answer is no; see the end of the next section.)

## 10 Examples of (infinite dimensional) subspaces

This section is a playful intermission: let's find (infinite-dimensional) spaces with (infinite-dimensional) subspaces.

**Example 10.1** Consider $\ell^\infty$. Elements of $\ell^\infty$ can be represented as bounded countable sequences $\{a_n\}_{n\in\mathbb{N}}$ with $a_n \in \mathbb{F}$. Let $e_n$ be the sequence with 1 in the $n$th spot and 0 elsewhere.

(1) $\{e_n : n \in \mathbb{N}\}$ is linearly independent in $\ell^\infty$.

(2) $\{e_n : n \in \mathbb{N}\}$ is not a basis of $\ell^\infty$.

(3) For all $p \geq 1$, $\ell^p$ is a subspace of $\ell^\infty$ that is not closed (in the $\ell^\infty$ norm): $\{\{1, 1/2^{1/p}, 1/3^{1/p}, \ldots, 1/n^{1/p}, 0, 0, \ldots\} : n \in \mathbb{N}\}$ is a sequence in $\ell^p$ that is Cauchy in the $\ell^\infty$ norm. The limit $\{1, 1/2^{1/p}, 1/3^{1/p}, \ldots, 1/n^{1/p}, \ldots\}$ is not in $\ell^p$.

(4) $c$, the set of all converging sequences, is a closed subspace of $\ell^\infty$; $c_0$, the set of all converging sequences converging to 0, is a closed subspace of $c$; the set of all sequences that are eventually 0 is a linear subspace of $c_0$ that is not closed.

(5) For $m \in \mathbb{N}$, $\{\{a_n\} \in \ell^\infty : a_{mn} = 0$ for all $n \in \mathbb{N}\}$ is a closed subspace of $\ell^\infty$.

(6) For $m \in \mathbb{N}$, $\{\{a_n\} \in \ell^\infty : a_n = m a_{mn}$ for all $n \in \mathbb{N}\}$ is a closed subspace of $\ell^\infty$.

(7) If $\mathbb{F} = \mathbb{R}$: For a strictly increasing sequence $\{m_n\}$ in $\mathbb{N}$ and a sequence $\{r_n\}$ in $\mathbb{R}$, $\{\{a_n\} \in \ell^\infty : a_{m_n} \geq r_n$ for all $n \in \mathbb{N}\}$ is a closed subset of $\ell^\infty$ that is not a subspace.
Alternatively, for arbitrary $\mathbb{F}$: For a strictly increasing sequence $\{m_n\}$ in $\mathbb{N}$ and a sequence $\{r_n\}$ in $\mathbb{R}$, $\{\{a_n\} \in \ell^\infty : |a_{m_n}| \geq r_n$ for all $n \in \mathbb{N}\}$ is not a subspace of $\ell^\infty$.
(The same conclusions if $\geq$ is replaced by $\leq$.)

(8) Let $\{m_n\}$ be a strictly increasing sequence in $\mathbb{N}$. Under what conditions on the sequence $\{r_n\}$ in $\mathbb{R}$, is $\{\{a_n\} \in \ell^\infty : a_{m_n} = r_n a_n$ for all $n \in \mathbb{N}\}$ a (closed) subspace of $\ell^\infty$?

(9) Let $M$ consist of all those sequences $\{a_n\}$ for which there exist $N \in \mathbb{N}$ and a polynomial $p(X)$ with coefficients in $\mathbb{F}$ such that for all $n \geq N$, $a_n = p(n)$. Well, since elements of $\ell^\infty$ are bounded, these $p$ must be constant polynomials, so all we are saying here is that $M$ is the set of all eventually constant sequences. Clearly $M$ is a linear subspace of $\ell^\infty$. It is not closed.

(10) Let $M'$ consist of all those sequences $\{a_n\}$ for which there exist $N \in \mathbb{N}$ and a polynomial $p(X)$ with coefficients in $\mathbb{F}$ such that for all $n \geq N$, $a_n = p(1/n)$. Since $p(1/n) + q(1/n) = (p+q)(1/n)$, $M'$ is a linear subspace of $\ell^\infty$. It is not closed (because the limit of polynomials need not be a polynomial).

(11) Let $\{m_n\}$ be a strictly increasing sequence in $\mathbb{N}$. What is $\{\{a_n\} \in \ell^\infty : \{a_{m_n}\}_n$ is a convergent sequence$\}$?

(12) Let $M = \{\{a_n\} \in \ell^\infty : a_{2n} = 0$ for all $n \in \mathbb{N}\}$ and $N = \{\{a_n\} \in \ell^\infty : n a_{2n} = a_{2n-1}$ for all $n \in \mathbb{N}\}$. Then $M$ and $N$ are closed linear subspaces of $\ell^\infty$. Let $\{x_n\}$ be an eventually zero sequence. Define $a_{2n} = 0$, $a_{2n-1} = x_{2n-1} - n x_{2n}$, $b_{2n} = x_{2n}$, $b_{2n-1} = n x_{2n}$. Since $x_n = 0$ for large $n$, it follows that $\{a_n\} \in M$ and $\{b_n\} \in N$. Clearly $\{a_n\} + \{b_n\} = \{x_n\}$. Thus $M + N$ contains all sequences that are eventually 0. Hence the closure of $M + N$ contains all converging sequences with limit 0. In particular, $\{1/\sqrt{n}\}$ is in the closure of $M + N$. Suppose that $\{1/\sqrt{n}\} \in M + N$.

Write $\{1/\sqrt{n}\} = \{a_n\} + \{b_n\}$ for some $\{a_n\} \in M$ and $\{b_n\} \in N$. Necessarily $b_{2n} = 1/\sqrt{2n}$, so that $b_{2n-1} = \sqrt{n}/\sqrt{2}$, but then $\{b_n\}$ is not in $\ell^\infty$, which gives a contradiction. It follows that the sum of two closed subspaces need not be closed.

The following is a modification of what we did in class. Take $p > 1$, and closed subspaces $M = \{\{a_n\} \in \ell^p : a_{2n} = 0 \text{ for all } n \in \mathbb{N}\}$ and $N = \{\{a_n\} \in \ell^p : na_{2n} = a_{2n-1} \text{ for all } n \in \mathbb{N}\}$ of $\ell^p$. Again all eventually zero sequences are in $M + N$, so that all elements in $\ell^p$ are in the closure of $M + N$. If $M + N$ is closed, then for any $r \in (1, p)$, it follows that $\{n^{-1/r}\} = \{a_n\} + \{b_n\}$ for some $\{a_n\} \in M$ and $\{b_n\} \in N$. Necessarily $b_{2n} = (2n)^{-1/r}$, so that $b_{2n-1} = 2^{-1/r}n^{1-1/r} = 2^{-1/r}n^{(r-1)/r}$, but then $\{b_n\}$ is not in $\ell^p$, which gives a contradiction. It follows that the sum of the closed subspaces $M$ and $N$ in $\ell^p$ is not closed.

## 11 Sublinear functional and the Hahn–Banach Theorem

**Definition 11.1** *Let $X$ be a vector space over $\mathbb{R}$. A function $p : X \to \mathbb{R}$ is a* **sublinear functional** *if for all $x, y \in X$ and all $r \in \mathbb{R}_{\geq 0}$, $p(rx) = rp(x)$ and $p(x + y) \leq p(x) + p(y)$.*

The main example of a sublinear functional is a norm, or a norm composed with a linear operator.

**Definition 11.2** *Let $X$ be a vector space over $\mathbb{F}$. A* **linear functional** *is a linear transformation $X \to \mathbb{F}$. (Explicitly, a function $p : X \to \mathbb{F}$ is a* **linear functional** *if for all $x, y \in X$ and all $r \in \mathbb{F}$, $p(rx) = rp(x)$ and $p(x + y) = p(x) + p(y)$.)*

Clearly all linear functionals to $\mathbb{F} = \mathbb{R}$ are sublinear, and some (but not all) sublinear functionals are linear.

**Theorem 11.3** (Hahn–Banach Theorem) *Let $X$ be a vector space over $\mathbb{R}$. Let $p : X \to \mathbb{R}$ be a sublinear functional. Let $M$ be a subspace of $X$ and $f_0$ a linear functional on $M$ such that for all $x \in M$, $f_0(x) \leq p(x)$. Then there exists a linear functional $f$ on $X$ such that*

*(1) $f|_M = f_0$.*

*(2) For all $x \in X$, $f(x) \leq p(x)$.*

*Proof.* If $X = M$, then $f = f_0$ works.

Consider the case where $X$ is spanned by $M$ and a vector $x \in X \setminus M$. Note that for all $\alpha \in \mathbb{R}$, the function $f : X \to \mathbb{R}$ defined by $f(rx + y) = r\alpha + f_0(y)$ is a well-defined linear functional on $X$. We need to find a suitable $\alpha \in \mathbb{R}$ such that for all $r \in \mathbb{R}$ and all $y \in M$, $r\alpha + f_0(y) \leq p(rx + y)$. This holds by definition for $r = 0$. If $r > 0$, the restriction is that $\alpha + f_0(\frac{1}{r}y) = \alpha + \frac{1}{r}f_0(y) = \frac{1}{r}(r\alpha + f_0(y)) \leq \frac{1}{r}p(rx + y) = p(x + \frac{1}{r}y)$, so that $\alpha \leq p(x + y_1) - f_0(y_1)$ for all $y_1 \in M$. If $r < 0$, the restriction is that $\alpha - f_0(\frac{1}{|r|}y) = \alpha + \frac{1}{r}f_0(y) = \frac{1}{r}(r\alpha + f_0(y)) \geq \frac{1}{r}p(rx + y) = -p(-x + \frac{1}{|r|}y)$, so that $\alpha \geq f_0(y_2) - p(-x + y_2)$ for all $y_2 \in M$.

But for all $y_1, y_2 \in M$, $f_0(y_1)+f_0(y_2) = f_0(y_1+y_2) \le p(y_1+y_2) = p(x+y_1-x+y_2) \le p(x+y_1)+p(-x+y_2)$, so that $f_0(y_2) - p(-x+y_2) \le p(x+y_1) - f_0(y_1)$. Thus $\alpha$ exists as we can choose

$$\alpha = \sup\{f_0(y_1) + f_0(y_2) : y_2 \in M\}.$$

Now consider the general case of $X$ and $M$. We will use Zorn's lemma. Let $\mathcal{P}$ be the set of all pairs $(M_1, f_1)$, where $M_1$ is a subspace of $X$ containing $M$ and $f_1$ is a linear functional on $M_1$ such that $f_1|_M = f_0$ and for all $x \in M_1$, $f_1(x) \le p(x)$. Then $\mathcal{P}$ is not empty as it contains $(M, f_0)$. We can impose a partial order $\le$ on $\mathcal{P}$: $(M_1, f_1) \le (M_2, f_2)$ if $M_1$ is a subspace of $M_2$ and $f_2|_{M_1} = f_1$. Let $L$ be a totally ordered subset of $\mathcal{P}$ (i.e., a chain). Write $L = \{(M_\alpha, f_\alpha) : \alpha \in I\}$ for some index set $I$. Let $\tilde{M} = \cup_\alpha M_\alpha$, and $\tilde{f} : \tilde{M} \to \mathbb{R}$ be defined by $\tilde{f}(v) = f_\alpha(v)$ whenever $v \in M_\alpha$. By the order on $\mathcal{P}$, this $\tilde{f}$ is well-defined, and it is a linear functional on $\tilde{M}$. Furthermore, by the definition, for all $x \in \tilde{M}$, $\tilde{f}(x) \le p(x)$. Thus $(\tilde{M}, \tilde{f})$ is an element of $\mathcal{P}$ that is an upper bound on $L$. Thus Zorn's lemma applies: there exists a maximal $(M', f')$ in $\mathcal{P}$. If $M' \ne X$, there exists $x \in X \setminus M'$. By the previous part, we can extend $f'$ to the strictly larger subspace spanned by $M'$ and $x$, which contradicts the maximality of $(M', f')$. Thus necessarily $X = M'$, and $f'$ is the desired $f$. □

**Theorem 11.4** (Hahn–Banach Theorem) *Let $X$ be a vector space over $\mathbb{C}$. Let $p : X \to \mathbb{R}$ be a sublinear functional. Let $M$ be a subspace of $X$ and $f_0$ a linear functional on $M$ such that for all $x \in M$, $Re(f_0(x)) \le p(x)$. Then there exists a linear functional $f$ on $X$ such that*

*(1) $f|_M = f_0$.*
*(2) For all $x \in X$, $Re(f(x)) \le p(x)$.*

*Proof.* Since $X$ is a vector space over $\mathbb{C}$, it is a vector space over $\mathbb{R}$. Thus by Theorem 11.3, there exists a linear functional $g : X \to \mathbb{R}$ such that $g|_M = Re(f_0)$ and for all $x \in X$, $g(x) \le p(x)$. Define $f : X \to \mathbb{C}$ as

$$f(x) = g(x) - ig(ix).$$

Clearly $g = Re(f)$, $f$ is $\mathbb{R}$-linear, and since $f(ix) = g(ix) - ig(i^2x) = g(ix) + ig(x) = if(x)$, it folows that $f$ is $\mathbb{C}$-linear as well. If $x \in M$, then $f(x) = g(x) - ig(ix) = Re(f_0(x)) - iRe(f_0(ix))) = Re(f_0(x)) - iRe(if_0(x))) = Re(f_0(x)) + iIm(f_0(x))) = f_0(x)$, which finishes the proof. □

**Remark 11.5** If $X$ is a normed vector space over $\mathbb{C}$, then it is a normed vector space over $\mathbb{R}$, and the function

$$\varphi : B_{\mathbb{C}}(X, \mathbb{C}) \to B_{\mathbb{R}}(X, \mathbb{R})$$

given by $\varphi(f) = Re(f)$ is an isometric isomorphism. (I leave the proof as an exercise.)

**Corollary 11.6** (Hahn–Banach Theorem) *Let $X$ be a vector space over $\mathbb{F}$. Let $p : X \to [0, \infty) \subseteq \mathbb{R}$ be a seminorm (or more generally(?), a sublinear functional with $p(rx) = |r|p(x)$ for all $r \in \mathbb{F}$ and $x \in X$). Let $M$ be a subspace of $X$ and $f_0$ a linear functional on $M$ such that for all $x \in M$, $|f_0(x)| \le p(x)$. Then there exists a linear functional $f$ on $X$ such that*

*(1) $f|_M = f_0$.*

*(2) For all $x \in X$, $|f(x)| \le p(x)$.*

*Proof.* If $\mathbb{F} = \mathbb{R}$, then for all $x \in M$, $f_0(x) \le |f_0(x)| \le p(x)$. Then by Theorem 11.3 there exists $f : X \to \mathbb{R}$ that extends $f_0$ and such that for all $x \in X$, $f(x) \le p(x)$. Hence $-f(x) = f(-x) \le p(-x) = p(x)$, so that $|f(x)| \le p(x)$.

Now suppose that $\mathbb{F} = \mathbb{C}$. Then for all $x \in M$, $Re(f_0(x)) \le |f_0(x)| \le p(x)$, so that by Theorem 11.4, there exists $f : X \to \mathbb{C}$ that extends $f_0$ and such that for all $x \in X$, $Re(f(x)) \le p(x)$. Now let $x \in X$ and $r, \theta \in [0, \infty)$ such that $f(x) = re^{i\theta}$. Hence $|f(x)| = r = e^{-i\theta}re^{i\theta} = e^{-i\theta}f(x) = f(e^{-i\theta}x) = Re(f(e^{-i\theta}x)) \le p(e^{-i\theta}x) = p(x)$. □

**Corollary 11.7** *Let $X$ be a normed vector space with a subspace $M$. If $f_0 \in M^*$, there exists $f \in X^*$ such that $f|_M = f_0$ and $\|f\| = \|f_0\|$.*

*Proof.* Let $p(x) = \|f_0\| \|x\|$. This is a sublinear functional, and for all $x \in M$, $|f_0(x)| \le \|f_0\| \|x\|$. Corollary 11.6 says that there exists $f : X \to \mathbb{F}$ such that $f|_M = f_0$ and $|f| \le p$. In other words, for all $x \in X$, $|f(x)| \le p(x) = \|f_0\| \|x\|$, which says that $\|f\| \le \|f_0\|$. The other inequality is clear. □

**Corollary 11.8** *If $X$ is a normed linear space and $x_0$ is non-zero in $X$, then there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x_0) = \|x_0\|$.*

*Proof.* Let $M$ be the span of $x_0$. Define $f_0(rx_0) = r\|x_0\|$ for all $r \in \mathbb{F}$. Then $f_0 \in M^*$, $\|f_0\| = 1$, and by the previous corollary, the $f$ with specified properties exists. □

**Corollary 11.9** *If $X$ is a normed vector space and $x_0 \in X$, then $f(x_0) = 0$ for all $f \in X^*$ implies that $x_0 = 0$.* □

**Corollary 11.10** *If $X$ is a normed linear space and $\{x_1, \ldots, x_n\}$ are linearly independent, then for arbitrary $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ there exists $f \in X^*$ such that for all $i = 1, \ldots, n$, $f(x_i) = \alpha_i$.*

*Proof.* Let $M$ be the span of $x_1, \ldots, x_n$. Define $f_0(\sum_i r_i x_i) = \sum_i r_i \alpha_i$ for all $r \in \mathbb{F}$. Then $f_0$ is linear, and since $M$ is finite-dimensional, $f_0$ is even continuous, so that $f_0 \in M^*$. By Corollary 11.9, the $f$ with specified properties exists. □

**Corollary 11.11** *Let $M$ be a closed subspace of a normed linear space $X$, and let $x_0 \in X \setminus M$. Then there exists $f \in X^*$ such that $f|_M = 0$, $\|f\| = 1$, and $f(x_0) =$ distance$(x_0, M)$.*

*Proof.* We apply Corollary 11.8 to the normed vector space $X/M$ and the non-zero element $x_0 + M$ in $X/M$. We know that distance$(x_0, M) = \|x_0 + M\|$. Thus there exists $f_1 \in (X/M)^*$ such that $\|f_1\| = 1$ and $f_1(x_0 + M) = \text{distance}(x_0, M)$. Define $f : X \to \mathbb{F}$ as the composition of the canonical map $X \to X/M$ and $f_0 : X/M \to \mathbb{F}$. Then for all $x \in M$, $f(x) = f_1(x+M) = f_1(0+M) = 0$. For all $x \in X$, $\|f(x)\| = \|f_1(x + M)\| \leq \|x + M\| \leq \|x\|$, so that $\|f\| \leq 1$. Fix $\epsilon > 0$. Then there exists $m \in M$ such that $\|x_0 + m\| < \|x_0 + M\| + \epsilon$, so that

$$f(x_0+m) = f_1(x_0+M) = \text{distance}(x_0, M) = \|x_0 + M\| \leq \|f\| \, \|x_0 + m\| \leq \|f\| \, (\|x_0 + M\|+\epsilon),$$

so that $\|f\| \geq \text{distance}(x_0, M)/(\|x_0 + M\| + \epsilon)$, which proves that $\|f\| \geq 1$, and hence that $\|f\| = 1$. □

**Theorem 11.12** *If $X$ is a normed vector space and $M$ is a linear subspace, then the topological closure $\overline{M}$ of $M$ equals*

$$\cap \,\text{kernel}\, f,$$

*where $f$ varies over elements of $X^*$ that have $M$ in the kernel. In particular, the topological closure is a vector subspace of $X$.*

*Proof.* By Exercise 4.3, each kernel $f$ is a closed subspace of $X$. Since it contains $M$, it contains $\overline{M}$. Thus $\cap$ kernel $f$ contains $\overline{M}$.

If $x_0 \notin \overline{M}$, then distance$(x_0, M) > 0$. Certainly distance$(x_0, \overline{M}) \leq$ distance$(x_0, M)$. For any $\epsilon > 0$, there exists $\overline{m} \in \overline{M}$ such that $d(x_0, \overline{m}) <$ distance$(x_0, \overline{M}) + \epsilon/2$, and then there exists $m \in M$ such that $\|m - \overline{m}\| < \epsilon/2$. Thus $\|x_0 - m\| \leq \|x_0 - \overline{m}\| + \|\overline{m} - m\| <$ distance$(x_0, \overline{M})+\epsilon$, so that distance$(x_0, M) \leq \epsilon+$distance$(x_0, \overline{M})$. As this holds for all $\epsilon$, we get that distance$(x_0, M) \leq$ distance$(x_0, \overline{M})$, whence distance$(x_0, M) =$ distance$(x_0, \overline{M})$. In particular, $x_0 \notin \overline{M}$.

By Corollary 11.11 there exists $f \in X^*$ such that for all $x \in \overline{M}$, $f(x) = 0$ and $f(x_0) =$ distance$(x_0, \overline{M}) =$ distance$(x_0, M)$. In particular, $x_0 \notin$ kernel $f$. □

**Corollary 11.13** *Let $X$ be a normed vector space and $M$ a linear subspace. Then $M$ is dense in $X$ if and only if any $f \in X^*$ that vanishes on $M$ is zero.* □

**Example 11.14** Let $\{e_1, e_2, \ldots\}$ be a basis of $\oplus_n \mathbb{F}$, and let $T : \oplus_n \mathbb{F} \to \mathbb{F}$ be the linear functional defined by $T(e_i) = i$. Then $T$ is not continuous, so $T \notin (\oplus_n \mathbb{F})^*$. Let $M$ be the kernel of $T$. The claim is that $M$ is not closed and that the closure of $M$ is the whole space $\oplus_n \mathbb{F}$. Namely, let $m \in M$. Define $x_i = m - \frac{T(m)}{i} e_i$. Then $T(x_i) = T(m) - \frac{T(m)}{i} T(e_i) = 0$, so that $x_i \in M$. The norm of $\frac{T(m)}{i} e_i$ is $\frac{|T(m)|}{i}$, which goes to 0 as $i$ goes to $\infty$, so that $\{x_i\}$ converges to $m$. This proves the claims.

**Theorem 11.15** *There exists $T \in (\ell^\infty)^*$ such that*

*(1) $\|T\| = 1$.*

*(2) If $x \in c$, then $T(x) = \lim x$.*

*(3) If $x \in \ell^\infty$ with $x_n \in [0, \infty)$ for all $n \in \mathbb{N}$, then $T(x) \in [0, \infty)$.*

*(4) If $x, y \in \ell^\infty$ and for all $n \in \mathbb{N}$, $y_n = x_{n+1}$, then $T(x) = T(y)$.*

*Proof.* First define $S : \ell^\infty \to \ell^\infty$ by $S(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$. Then $S$ is a linear function, and $\|S(x)\| \leq \|x\|$, so that $S$ is continuous. Condition (4) says that for all $x \in \ell^\infty$, $T \circ S(x) = T(x)$.

Let $M = \{x - S(x) : x \in \ell^\infty\}$. Then $M$ is a vector subspace of $X$.

Let $y = \{1, 1, 1, \ldots\} \in \ell^\infty$. Since $0 \in M$, distance$(y, M) \leq 1$. Suppose that there exist $x \in \ell^\infty$ and $\alpha \in (0, 1)$ such that for all $n \geq 1$, $|x_n - x_{n-1} - 1| < \alpha$ (implicitly, $x_0 = 0$). Note that $Re(x) \in M$, and the assumption above says that for all $n$,

$$-\alpha < Re(x_n) - Re(x_{n-1}) - 1 < \alpha.$$

In particular, $1 - \alpha < Re(x_1) < 1 + \alpha$, and for all $n > 1$, $1 - \alpha + Re(x_{n-1}) < Re(x_n) < 1 + \alpha$. It follows that $\{Re(x_n)\}$ is a bounded strictly increasing sequence of positive real numbers, so that $\lim_n Re(x_n)$ exists and is a real number. But then $1 = \lim_n |Re(x_n) - Re(x_{n-1}) - 1| \leq \alpha < 1$ gives a contradiction. Thus no such $\alpha$ exists, so that for all $x \in \ell^\infty$, distance$(y, x) = \sup |x_n - x_{n-1} - 1| \geq 1$.

By triangle inequality it follows that distance$(y, \overline{M}) = 1$. (And the topological closure $\overline{M}$ is still a vector subspace.)

Thus by Corollary 11.11, there exists $T \in (\ell^\infty)^*$ such that $T|_{\overline{M}} = 0$, $\|T\| = 1$, and $T(y) = \text{distance}(y, M) = 1$.

Note that $T$ satisfies the desired properties (1) and (4).

Let $x \in \oplus_n \mathbb{F}$, i.e., $x$ is a finite sequence. Then for some $n \in \mathbb{N}$, $S^n(x)$ is the zero sequence, and in particular, $S^n(x)$ is in the kernel of $T$. Since for all $m$, $S^m(x) - S^{m+1}(x) \in M \subseteq \text{kernel}\, T$, it follows that $x = \sum_{m=1}^{n} (S^m(x) - S^{m+1}(x)) \in \text{kernel}\, T$. Thus $c_0(\mathbb{N}) \in \text{kernel}\, T$.

Let $x \in c^*$ be a sequence with limit 0. Then $x$ is in the closure of $\oplus_n \mathbb{F}$. Since $T$ is continuous, it follows that $x \in \text{kernel}\, T$. Thus $c_0 \in \text{kernel}\, T$.

The desired condition (2) follows: for any $x \in c$, if $r$ is the limit, then $x - ry \in c_0 \subseteq \text{kernel}\, T$, whence $T(x) = T(x - ry + ry) = T(x - ry) + rT(y) = 0 + r \cdot 1 = r = \lim x$.

Finally we prove Condition (3). Let $x \in \ell^\infty$ with $x_n \in [0, \infty)$ for all $n$. If $x_n = 0$ for all $n$, then $T(x) = 0 \in [0, \infty)$. So we may assume that $x \neq 0$. Then $\|x\| > 0$, and for all $n$, $\frac{x_n}{\|x\|} \in [0, 1]$. Then $\left\| y - \frac{x}{\|x\|} \right\| \leq 1$, and $|1 - T(\frac{x}{\|x\|})| = |T(y - \frac{x}{\|x\|})| \leq \|T\| \left\| y - \frac{x}{\|x\|} \right\| \leq 1$. It follows that $1 - Re(T(\frac{x}{\|x\|})) \leq |1 - T(\frac{x}{\|x\|})| \leq 1$, whence $Re(T(\frac{x}{\|x\|})) \geq 0$, so that $Re(T(x)) \geq 0$.

* Recall: $c$ denotes the set of all convergent sequences, and $c_0$ denotes the set of all convergent sequences with limit 0.

In case $\mathbb{F} = \mathbb{R}$, we are done with the proof. Otherwise, $\mathbb{F} = \mathbb{C}$, and we can go back and construct $T$ as above on $\ell^\infty(\mathbb{R}) \to \mathbb{R}$. Set $T' : \ell^\infty(\mathbb{C}) \to \mathbb{C}$ with $T'(x) = T(Re(x)) + iT(Im(x))$. This is an additive and $\mathbb{R}$-linear function, and since $T'(ix) = T(Re(ix)) + iT(Im(ix)) = T(-Im(x)) + iT(Re(x)) = -T(Im(x)) + iT(Re(x)) = iT'(x)$, it follows that $T'$ is $\mathbb{C}$-linear. For all $x \in \ell^\infty$, $|T'(x)|^2 = |T(Re(x)) + iT(Im(x))|^2 = (T(Re(x)))^2 + (T(Im(x)))^2 \leq \|T\|^2 ((Re(x))^2 + (Im(x))^2) = \|T\|^2 \|x\|^2 = \|x\|^2$, so that $\|T'\| \leq \|T\| = 1$. However, since $T'$ extends $T$, $\|T'\| \geq \|T\|$, so that $\|T'\| = 1$. Thus condition (1) holds for $T'$, and trivially the other conditions hold for $T'$ as well. □

The linear functional from Theorem 11.15 is called a **Banach limit**.

Banach limits do not behave like **limits** in the following sense: there exist $x, y \in \ell^\infty$ such that $T(xy) \neq T(x)T(y)$. Namely, let $x = (2, 17, 2, 17, 2, \ldots)$ and $y = (17, 2, 17, 2, \ldots)$. By Condition (4) of Banach limits, $T(x) = T(y)$. Then $2T(x) = T(x) + T(y) = T(x + y) = T(19, 19, 19, \ldots) = 19$ by Condition (2) of Banach limits, so that $T(x) = 9.5$. Similarly, since $xy = \{34\}$, $T(xy) = 34$. Hence $T(xy) \neq T(x)T(y)$.

What falls out of the calculation above: if $x$ is a periodic sequence of period $n$, then $T(x)$ is the average $\frac{x_1 + \cdots + x_n}{n}$. Thus $T$ is uniquely determined for periodic or eventually periodic sequences.

**Question 11.16** (This was raised in class. I don't know an answer.) What is $T(\{3, 1, 4, 1, 5, 9, 2, 6, \ldots\})$?

Banach limits are not unique, in fact, so the question above probably has multiple answers. According to Larry Brown, they correspond to probability measures on the corona of $\mathbb{N}$ (corona of a completely regular topological space $X$ is the complement of $X$ in its Stone-Čech compactification).

## 12 A section of big theorems

**Definition 12.1** *A subset $E$ of a metric space $X$ is* **nowhere dense** *if the complement of the closure of $E$ is dense.*

In other words, $E$ is nowhere dense if and only if $\overline{E}$ has empty interior.

For example, $\mathbb{N}$ is nowhere dense in $\mathbb{R}$. Some countable subsets of $\mathbb{R}$, such as $\mathbb{Q}$, are not nowhere dense in $\mathbb{R}$. The Cantor set is nowhere dense in $[0, 1]$.

**Lemma 12.2** *Let $X$ be a metric space. For each $x \in X$ and $r > 0$ there exists $s > 0$ such that $\overline{B(x, s)} \subseteq B(x, r)$. In fact, any $s < r$ works.*

*Proof.* Let $0 < s < r$. Let $y \in \overline{B(x, s)}$. Then for all $n \in \mathbb{N}$ there exists $y_n \in B(x, s)$ such that $d(y, y_n) < 2^{-n}$. In particular, for $n \in \mathbb{N}$ such that $r - s > 2^{-n}$,

$$d(y, x) \leq d(y, y_n) + d(x, y_n) < 2^{-n} + s < r,$$

so that $\overline{B(x, s)} \subseteq B(x, r)$. □

**Theorem 12.3** (Baire Category Theorem) *A non-empty complete metric space is not the union of a countable collection of nowhere dense sets.*

*Proof.* Let $E_1, E_2, \ldots$ be nowhere dense sets in the metric space $X$ such that $\cup_n E_n = X$. Then $O_n = X \setminus \overline{E}_n$ is dense and open in $X$, and $\cap O_n = \emptyset$.

We will prove that countable intersections of dense open sets cannot be empty. Since $X$ is non-empty and $O_1$ is dense in $X$, necessarily $O_1 \neq \emptyset$. Let $x_1 \in O_1$. Since $O_1$ is open, there exists $r_1 > 0$ such that $B(x_1; r_1) \subseteq O_1$. Since $O_2$ is dense, $O_2 \cap B(x_1; r_1) \neq \emptyset$, so choose $x_2 \in O_2 \cap B(x_1; r_1)$. Since $O_2 \cap B(x_1; r_1)$ is open, there exists $r_2 > 0$ such that $B(x_2; r_2) \subseteq O_2 \cap B(x_1; r_1)$. By possibly replacing $r_2$ even by a smaller number, we may assume that $r_2 \leq r_1/2$ and that $\overline{B(x_2; r_2)} \subseteq O_2 \cap B(x_1; r_1)$. In general, similarly there exist $x_n \in X$ and $r_n > 0$ such that $r_n \leq r_1/2^n$ and $\overline{B(x_n; r_n)} \subseteq O_n \cap B(x_{n-1}; r_{n-1})$. Thus $\{x_n\}$ is a Cauchy sequence in $X$, and since $X$ is complete, there exists $x \in X$ such that $x_n \to x$. Since for all $m \geq n$, $x_m \in \overline{B(x_n; r_n)}$, it follows that $x \in \overline{B(x_n; r_n)}$ for all $n$, whence $x \in O_n$ for all $n$. □

The complete assumption is necessary: $\mathbb{Q}$ is a countable collection of singleton sets, thus of nowhere dense sets.

**Definition 12.4** *A function is* **open** *if it the image of every open set is open. A linear function $T$ is* **almost open** *if for all $r > 0$ there exists $s > 0$ such that the closure of $T(B(0, r))$ contains $B(0, s)$.*

Clearly an open linear function is almost open and surjective. However, not all almost open linear functions are open. For example, the inclusion $\oplus_{n \in \mathbb{N}} \mathbb{F} \to \ell^1$ is almost open, but is not open. Note that the inclusion $\oplus_{n \in \mathbb{N}} \mathbb{F} \to \ell^\infty$ is not almost open. Not all surjective linear functions are almost open: for example, the map $T : X = \oplus_n \mathbb{F} \to Y = \oplus_n \mathbb{F}$ given by $T(e_n) = e_n/n$ is surjective, continuous, and not almost open. Namely, since $\| \ \|$ in $X$ and $Y$ is given by $\|\sum_n c_n e_n\| = \sum_n |c_n|$, $\|T(\sum_n c_n e_n)\| = \|\sum_n c_n/n e_n\| = \sum_n |c_n/n| \leq \sum_n |c_n| = \|\sum_n c_n e_n\|$, it follows that $T$ is continuous and that $\|T\| \leq 1$. Since $T(e_1) = e_1$, it follows that $\|T\| = 1$. If for some positive $s$, $B_Y(0, s) \subseteq \overline{T(B_X(0, 1))}$, then for all $m$, $s e_m/2 \in B_Y(0, s)$ can be approximated arbitrarily closely by $T(x)$ for $x \in B(0, 1)$. Write $x = \sum_n c_n e_n$. Necessarily $c_m/m$ has to be arbitrarily close to $s/2$, but then $c_m$ must be close to $sm/2$, but then $x = \sum_n c_n e_n$ cannot have norm at most 1. Thus $T$ is not almost open.

**Remark 12.5** Let $M$ be a closed linear subspace of a normed vector space $X$. Define $\pi : X \to X/M$ by $\pi(x) = x + M$. We know that $\pi$ is linear. Since $\|\pi(x)\| = \|x + M\| \leq \|x\|$, we have that $\pi$ is continuous and $\|\pi\| \leq 1$. This map is even open! Namely, suppose that $\|y + M\| < 1$. Then there exists $m \in M$ such that $\|y + m\| < 1$. But $\pi(y + m) = y + M$, so that $B(0, 1) \subseteq \pi(B(0, 1))$. Also, if $\|x\| < 1$, then $\|\pi x\| < 1$. This proves that $\pi(B(0, 1)) \subseteq B(0, 1)$. It follows that $B(0, 1) = \pi(B(0, 1))$, and by linearity that $B(0, r) = \pi(B(0, r))$ for all $r > 0$.

A consequence of the remark is that for $U \subseteq X/M$, $U$ is open if and only if $\pi^{-1}(U)$ is open.

And yet another consequence is the *Universal Mapping Theorem of quotients*: If $Y$ is a topological space, a function $f : X/M \to Y$ is continuous if and only if $f \circ \pi$ is continuous.

In particular, if $T \in B(X, Y)$, where $X$ and $Y$ are normed vector spaces, and if $M$ is a closed subpace of $X$ contained in the kernel of $T$, then there exists a unique $\tilde{T} : B(X/M, Y)$ such that $T = \tilde{T} \circ \pi$. There is a natural one-to-one correspondence between the sets $\{T \in B(X, Y) : T|_M = 0\}$ and $B(X/M, Y)$.

For any $y + M$ with $\|y + M\| < 1$, there exists $x \in y + M$ such that $\|x\| < 1$. Thus $\tilde{T}(y + M) = T(x) \leq \|T\| \|x\| \leq \|T\|$, so that by Proposition 8.6, $\left\|\tilde{T}\right\| \leq \|T\|$. Also, $\|T\| = \left\|\tilde{T} \circ \pi\right\| \leq \left\|\tilde{T}\right\| \|\pi\| \leq \left\|\tilde{T}\right\|$, whence $\left\|\tilde{T}\right\| = \|T\|$.

Note that if $M \neq X$, non-zero $T$ and $\tilde{T}$ exist by the Hahn–Banach Theorem (Corollary 11.11), so that all this forces $\|\pi\| = 1$.

**Theorem 12.6** (The Banach Open Mapping Theorem) *Let $X$ and $Y$ be normed vector spaces, with $X$ Banach. Then any almost open element of $B(X, Y)$ is open.*

*Proof.* Let $T \in B(X, Y)$ be almost open. Without loss of generality $Y \neq 0$, and then necessarily $T \neq 0$. By assumption there exists $s > 0$ such that $B(0, s) \subseteq \overline{T(B(0, 1))}$. Since $T$ is linear, it follows that for all $r > 0$, $B(0, rs) \subseteq \overline{T(B(0, r))}$.

We will prove that $B(0, rs/3) \subseteq T(B(0, r))$.

Let $y \in Y \setminus \{0\}$. Then for every $\alpha > 1$, in particular for $\alpha = 1 + s/(4\|T\|)$, $y \in B(0, \alpha\|y\|) \subseteq \overline{T(B(0, \alpha\|y\|/s)}$, so that there exists $x \in B(0, \alpha\|y\|/s)$, such that $\|y - T(x)\| < \|y\|/4$. Then $x/\alpha \in B(0, \|y\|/s)$, and $\|y - T(x/\alpha)\| \leq \|y - T(x)\| + \|T(x) - T(x/\alpha)\| < \|y\|/4 + \|T\|\|x\|(1 - 1/\alpha) < \|y\|/4 + \|T\|\alpha\|y\|(1 - 1/\alpha)/s = \|y\|/4 + \|T\|\|y\|(\alpha - 1)/s \leq \|y\|/2$. What this proves is that for every non-zero $y \in Y$, there exists $x_1 \in B(0, \|y\|/s)$ such that $\|y - T(x_1)\| < \|y\|/2$.

Set $y_1 = y - T(x_1)$. If $y_1 = 0$, we are done, so we may assume that $y_1 \neq 0$. Now construct $x_2$ from $y_1$ in the same way that $x_1$ was constructed from $y$, etc. Summary for $n \geq 1$ if $(y_{n-1} \neq 0)$:

$$\begin{aligned} y_n &= y - Tx_1 - \cdots - Tx_n, \\ \|y_n\| &= \|y - Tx_1 - \cdots - Tx_n\| < 2^{-n}\|y\|, \\ \|x_n\| &\leq \frac{1}{s}\|y_{n-1}\| < 2^{-n+1}\frac{1}{s}\|y\|. \end{aligned}$$

Since $\sum \|x_n\| < \infty$ and $X$ is complete, by Homework 4.1, $\sum x_n$ converges. Let $x = \sum x_n$. Since $T$ is continuous, $y - T(x) = \lim(y - Tx_1 - \cdots - Tx_n) = \lim y_n = 0$. Also, $\|x\| \leq \sum \|x_n\| \leq \frac{1}{s}\|y\| \sum_n 2^{-n+1} = \frac{2}{s}\|y\|$. This proves that $B(0, rs/3) \subseteq T(B(0, r))$.

Now let $U$ be an open subset of $X$, and let $x \in U$. Since $U$ is open, there exists $r > 0$ such that $B(x, r) \subseteq U$. In other words, $x + B(0, r) \subseteq U$. Then $T(x) + T(B(0, r)) = T(x +$

$B(0,r)) \subseteq T(U)$, and by the claim above, $B(T(x), rs/3)) = T(x) + B(0, rs/3)) \subseteq T(U)$, which proves that $T(U)$ is open. □

The complete assumption on $X$ is necessary: the inclusion $\oplus_{n\in\mathbb{N}}\mathbb{F} \to \ell^1$ is almost open, but not open. The almost open assumption is of course also necessary (say take the zero map into a non-trivial normed vector space).

**Theorem 12.7** (The Open Mapping Theorem) *Let $X$ and $Y$ be Banach spaces, and let $T \in B(X, Y)$ be surjective. Then $T$ is an open mapping.*

*Proof.* For each $n \in \mathbb{N}$, let $A_n = T(B(0, n))$. Since $X = \cup_n B(0, n)$ and $T$ is surjective, $Y = \cup_n A_n$. By the Baire Category Theorem (Theorem 12.3), since $Y$ is complete, there exists $N \in \mathbb{N}$ such that $A_N$ is not nowhere dense. This means that $Y \setminus \overline{A_N}$ is not dense, so there exists $y_0 \in \overline{A_N}$ such that for some $m > 0$, $B(y_0, m) \subseteq \overline{A_N} = \overline{T(B(0, N))}$.

We will prove that $B(0, m) \subseteq \overline{T(B(0, 2N))}$. Let $y \in B(0, m) \subseteq Y$. Then $y_0, y + y_0 \in B(y_0, m) \subseteq \overline{T(B(0, N))}$. Thus there exist sequences $x_n, x'_n \in B(0, N)$ such that $T(x_n) \to y_0$ and $T(x'_n) \to y+y_0$. Thus $T(x'_n - x_n) \to y$, and $x'_n - x_n \in B(0, 2N)$. Thus $y \in \overline{T(B(0, 2N))}$, and since $y$ was arbitrary in $B(0, m)$, we get that $B(0, m) \subseteq \overline{T(B(0, 2N))}$.

Since $T$ is linear, it follows that for all $r > 0$, $B(0, rm/(2N)) \subseteq \overline{T(B(0, r))}$. Thus $T$ is almost open, and so by Theorem 12.6, $T$ is open. □

**Examples 12.8** There are five assumptions in the theorem above: $X$ complete, $Y$ complete, $T$ linear, $T$ continuous, $T$ surjective. We show next which of these assumptions can be omitted.

(1) $X$ has to be complete. Otherwise, here is a counterexample. Let $e_n$ be a sequence in $c_0$ that has 1 in the $n$th spot and 0 elsewhere. We may also pick $b_0 = \{2^{-n}\}$. Then $\{b_0, e_1, e_2, \ldots\}$ is a linearly independent subset of $c_0$, so by Zorn's lemma there exists a basis $B$ of $c_0$ that contains $\{b_0, e_1, e_2, \ldots\}$. Or we could pick an arbitrary basis not containing $\{e_1, e_2, \ldots\}$, and then some $b_0 \in B \setminus \{e_1, e_2, \ldots\}$. Let $X'$ be the span of $B \setminus \{b_0\}$. Since $c_0$ is the completion of $Span\{e_1, e_2, \ldots\}$, $X'$ is not complete. Set $Y = c_0/Span(b_0)$. Since $Span(b_0)$ is a finite-dimensional subspace of $c_0$, it is closed in $c_0$, so that by Theorem 8.14, $Y$ is a Banach space. Define $T' : X' \to Y$ by $T'(e_n) = e_n/n$ for all $n$, $T'(b) = b$ for all $b \in B \setminus \{b_0, e_1, e_2, \ldots\}$, and extend $T'$ linearly to all of $X'$. Then $T'$ is linear, surjective, and even injective! By Exercise 7.4, it is likely that $T'$ is not continuous. Let $X = \{(x, T'(x)) : x \in X'\}$. Then $X$ is a linear space with norm $\|(x, T'(x))\| = \|x'\| + \|T'(x)\|$ (recall that the norm in $c_0$ is the sup norm). Since $X'$ is not complete, neither is $X$, but actually this may be hard to see directly without knowing more about the missing element $b_0$ from $X'$: namely, it may be that for a sequence $\{x_n\}$ in $X'$ converging to $b_0 \in c_0$, $\{(x_n, T'(x_n))\}$ is not Cauchy in $X$ and that may be hard to analyze without knowing more about $b_0$. Let's postpone this issue for a bit. Finally define $T : X \to Y$ by $T(x, T'(x)) = T'(x)$. Then clearly $T$ is surjective, injective, linear, and it is even continuous: $\|T(x, T'(x))\| = \|T'(x)\| \leq \|(x, T'(x))\|$. However, $T$ is

not open, as we show next. Let $s > 0$, and let $n \in \mathbb{N}$ with $1/n < s$. Note that $e_n/n \in B_Y(0,s)$. Let $x' \in X'$ such that $T'(x') = e_n/n$. Since $T'$ is injective, $x' = e_n$, so that $(x, T'(x'))$ in $X$ has norm at least 1, so that $B_Y(0,x) \not\subseteq T(B_X(0,1))$. Thus $T$ is not open. Now that we know that $T$ is not open, by Theorem 12.7, we can conclude definitively that $X$ cannot be complete.

Purely algebraically, $X = X/\operatorname{kernel} T \cong ImT = Y$ as vector spaces, but since $T$ does NOT preserve the norm, this is NOT an isomorphism of normed vector spaces, that is how one isomorphic copy of the vector space can be complete and the other not.

(2) $Y$ needs to be assumed complete. One might be tempted to think that the image of a Banach space by a continuous linear function would also be complete, but that is not the case. Let $X = \ell^1$, and let $Y = \{\{\frac{x_n}{n}\} : \{x_n\} \in \ell^1\}$. Then $Y$ is a linear subspace of $\ell^1$, but it is not closed. Namely $\{\frac{1}{n^2}\} \in \ell^1$ is not in $Y$, but it is the limit in the $\ell^1$-norm of the sequence $\{y_n\}$ in $Y$, where $y_n = \{1, 1/2^2, 1/3^2, \ldots, 1/n^2, 0, 0, \ldots\}$.

With these same $X, Y$, let $T : X \to Y$ be defined by $T(\{x_n\}) = \{\frac{x_n}{n}\}$. This $T$ is linear, surjective, injective, continuous, but it is not open. For all $s > 0$, pick $n$ such that $1/n < s$. Let $e_n \in \ell^1$ be 1 in the $n$th position and 0 elsewhere. Then $e_n/n$ in $Y$ has norm strictly smaller than 1. Then $\frac{e_n}{n} \in B_Y(0,s)$, but $e_n \notin B_X(0,1)$.

Again, purely algebraically, $X = X/\operatorname{kernel} T \cong ImT = Y$ as vector spaces, but since $T$ does NOT preserve the norm, this is NOT an isomorphism of normed vector spaces, and that is how one isomorphic copy of the vector space can be complete and the other not.

(3) $T$ needs to be linear, for otherwise take $X = \mathbb{R}^2$, $Y = \mathbb{R}$, $T(x,y) = xy^2$. Then $T$ is surjective, continuous, not linear, and not open because $T$ takes the open box $(0,1) \times (-1,1)$ onto non-open $[0,1)$.

(4) $T$ needs to be continuous, for otherwise we have the following counterexample. Let $X = \ell^1$. Let $e_n \in X$ have 1 in the $n$th spot and 0 elsewhere. Then $\{e_1, e_2, \ldots\}$ is a linearly independent subset of $X$, each of which has norm 1. By Zorn's lemma we may extend this linearly independent set to a basis $B$ of $X$. By Cauchy-Schwartz we may assume that each element of $B$ has norm 1. Set $Y = X$, and define $T : X \to Y$ by $T(e_n) = n^{(-1)^n} e_n$, and $T(b) = b$ for all $b \in B \setminus \{e_1, e_2, \ldots\}$. Extend $T$ to all of $X$ by linearity. Then $T$ is linear and bijective. It is not continuous because $T(e_{2n}) = 2ne_{2n}$ has norm $2n$. If there exists $s > 0$ such that $B_Y(0,s) \subseteq T(B_X(0,1))$, then in particular for $n > s$, $\frac{1}{2n-1}e_{2n-1} \in B_Y(0,s) \subseteq T(B_X(0,1))$, but $\frac{1}{2n-1}e_{2n-1}$ equals the image by $T$ of $e_{2n-1}$ only, and $e_{2n-1}$ is not in $B_X(0,1)$.

(5) Since every open map is surjective, necessarily the assumption $T$ being surjective must stay, and it is not implied by the other four assumptions (say take the zero map and $Y$ non-zero).

**Theorem 12.9** (The Inverse Function Theorem) *If $X$ and $Y$ are Banach spaces, $T \in B(X,Y)$,* kernel $T = \{0\}$ *and $T$ is surjective, then the inverse of $T$ exists and is continuous.*

*Proof.* By assumptions that kernel $T = \{0\}$ and that $T$ is surjective, $T$ is bijective. By Theorem 12.7, $T$ is open. Thus for all $r > 0$ there exists $s > 0$ such that $B(0,s) \subseteq T(B(0,r)) = (T^{-1})^{-1}(B(0,r))$, which says that so that $T^{-1}$ is continuous at 0, whence it is continuous. □

**Examples 12.10** There are six assumptions in the theorem above: $X$ complete, $Y$ complete, $T$ linear, $T$ continuous, $T$ injective, $T$ surjective. We show next that none of these assumptions can be omitted. First $X$ and $Y$ need to be complete, and $T$ needs to be continuous and surjective, by using the corresponding examples in Examples 12.8. Furthermore, $T$ needs to be injective for the inverse to exist. In addition, $T$ needs to be linear, for otherwise we have the following counterexample: Let $X = Y = \ell^\infty$ over $\mathbb{R}$, and let $T : X \to Y$ be given by $T(x_n) = (x_1, x_2 + (|x_1|+1)^{1/2}, x_3 + (|x_1|+1)^{1/3}, x_4 + (|x_1|+1)^{1/4}, \ldots)$. Clearly $T$ is surjective, injective, and not linear. It is continuous at every $c \in X$: Let $\epsilon > 0$. Let $x \in X$ satisfy $\|x - c\|_\infty < \min\{\epsilon/(\|c\|_\infty + 2), 1\}$ and in addition if $c_1 \neq 0$, then $x \in X$ satisfies $\|x - c\|_\infty < |c_1|/2$. This guarantees that $x_1$ and $c_1$ have the same sign (or $c_1 = 0$). Then for all $n > 1$,

$$\left|\left(x_n + (|x_1|+1)^{1/n}\right) - \left(c_n + (|c_1|+1)^{1/n}\right)\right| \le |x_n - c_n| + \left|(|x_1|+1)^{1/n} - (|c_1|+1)^{1/n}\right|$$

It suffices to prove that $\left|(|x_1|+1)^{1/n} - (|c_1|+1)^{1/n}\right|$ is less than $\epsilon/2$. Set

$$S = (|x_1|+1)^{(n-1)/n} + (|x_1|+1)^{(n-2)/n}(|c_1|+1)^{1/n} + \cdots + (|c_1|+1)^{(n-1)/n}.$$

This $S \ge 1$, and $S((|x_1|+1)^{1/n} - (|c_1|+1)^{1/n}) = |(|x_1| - |c_1|)|$, so that for all $n > 1$,

$$|(T(x) - T(c))_n| < \epsilon/2 + \left|(|x_1|+1)^{1/n} - (|c_1|+1)^{1/n}\right| \frac{S}{S} < \epsilon.$$

This proves that $T$ is continuous. However, $T^{-1}$ is not continuous. Namely, if $c = (0,0,0,\ldots)$, then $T^{-1}(c) = (0,0,\ldots)$, and for all $m,n > 1$, $(T^{-1}(1/m,0,0,\ldots))_n = (1/m, -1/(m+1)^{1/2}, -1/(m+1)^{1/3}, -1/(m+1)^{1/4}, \ldots)$, so that $\left\|T^{-1}(1/m,0,0,\ldots)\right\|_\infty = 1$ for all $m$. Thus $T^{-1}$ is not continuous at 0.

**Definition 12.11** *For any function $f : X \to Y$, the* **graph of f** *is*

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

**Theorem 12.12** (The Closed Graph Theorem) *If $X$ and $Y$ are Banach spaces and $T : X \to Y$ is linear, then $T$ is continuous if and only if the graph of $T$ is closed in $X \oplus Y$ (which norm on the direct product?).*

*Proof.* Suppose that $T$ is continuous. Let $\{x_n, y_n\} \in \Gamma(T)$ converge to $(x, y) \in X \oplus Y$. Then $x_n \to x$ and $y_n \to y$, and since $T$ is continuous, $y_n = T(x_n)$ converges to $y$ and $T(x)$, so that $y = T(x)$, and $(x, y) \in \Gamma(T)$.

Now suppose that $\Gamma(T)$ is closed. Since it is a linear subspace of $X \oplus Y$, it is a Banach space. The linear homomorphism $\pi_X : X \oplus Y \to X$ given by $\pi_X(x, y) = x$, is bounded as for all $(x, y)$, $\|\pi_X(x, y)\| = \|x\| \leq \|(x, y)\|$. Then $\pi = \pi_X|_{\Gamma(T)}$ is continuous, surjective, and injective. Thus by Theorem 12.9, $\pi^{-1}$ is continuous.

Now let $\pi_Y : X \oplus Y \to Y$ be given by $\pi_Y(x, y) = y$. Also this is linear and continuous. (It need not be injective even when restricted to $\Gamma(T)$.)

Note that $T = \pi_Y \circ \pi^{-1}$. Then $T$ is a composition of continuous linear functions, so $T$ is continuous. □

Rephrasing of the closedness of the graph: $\Gamma(T)$ is closed if and only if for all $x_n$ in $X$ with $x_n \to x$ and $T(x_n) \to y$, necessarily $y = T(x)$. Since $T$ is linear, $\Gamma(T)$ is closed if and only if for all $x_n$ in $X$ with $x_n \to 0$ and $T(x_n) \to y$, necessarily $y = 0$.

**Remark 12.13** Consider a measure space $(X, \Sigma, \mu)$. Suppose that for some $p, q \geq 1$, $\mathbb{L}^p(X, \Sigma, \mu) \subseteq \mathbb{L}^q(X, \Sigma, \mu)$. Let $T : \mathbb{L}^p \to \mathbb{L}^q$ be the inclusion. Then $\Gamma(T)$ is closed: Suppose that $f_n \to 0$ in $\mathbb{L}^p$ and $T(f_n) \to g$ in $\mathbb{L}^q$. Since $f_n \to 0$ in the norm, for every $m > 0$ there exists $N_m$ such that for all $n \geq N_m$, $\|f_n\| < 2^{-m}$. We may assume that $N_1 < N_2 < \cdots$. For $m \leq k$ let $S_{mk} = \{x \in X : |f_{N_k}(x)| \geq 2^{-m/2}\}$. Then $S_{mk} \in \Sigma$, and

$$2^{-kp} > \int |f_{N_k}|^p d\mu \geq \int_{S_{mk}} |f_{N_k}|^p d\mu \geq \mu(S_{mk}) 2^{-mp/2},$$

so that $\mu(S_{mk}) \leq 2^{-p(k-m/2)}$. Then $S_m = \cup_{k \geq m} S_{mk}$ has measure at most $2^{-pm/2+1}$. Then $\cap_{m \geq m_0} S_m$ has measure 0 for all $m_0$, so that $S = \cup_{m_0} \cap_{m \geq m_0} S_m$ has measure 0. Now let $x \in X \setminus S$. Since $x \notin S$, there exist infinitely many $m$ such that $x \notin S_m$. Thus for each of these $m$ and each $k \geq m$, $x \notin S_{mk}$. In particular, $|f_{N_k}(x)| < 2^{-m/2}$ for all $k \geq m$ and for infinitely many $m$. Thus $f_{N_k}(x) \to 0$ pointwise almost everywhere. Then also $T(f_{N_k})(x) \to 0$ pointwise almost everywhere, so $g = 0$ a.e., so $g = 0$ in $\mathbb{L}^q$. Thus we can conclude that $T$ is continuous, so there exists $c > 0$ such that for all $f \in \mathbb{L}^p$, $\|f\|_q \leq c \|f\|_p$.

**Theorem 12.14** (The Uniform Boundedness Principle) *Let $X$ be a Banach space and $Y_\alpha$ a normed linear space as $\alpha$ varies over some index set $I$. Let $T_\alpha \in B(X, Y_\alpha)$, and assume that for all $x \in X$, $\{\|T_\alpha(x)\| : \alpha\}$ is bounded. Then $\{\|T_\alpha\| : \alpha\}$ is bounded.*

The proof in the book has all $Y_\alpha$ the same.

*Proof.* Let $A_n = \{x \in X : \|T_\alpha(x)\| \leq n \text{ for all } \alpha\}$. Then $A_n$ is a closed set! Also, $\cup A_n = X$. By the Baire Category Theorem (Theorem 12.3), some $A_{n_0}$ has interior, i.e., there exist $x_0 \in X$ and $r > 0$ such that $B(x_0, r) \subseteq A_{n_0}$.

If $\|x\| < r$, then $x_0, x_0 + x \in A_{n_0}$. Then for all $\alpha$, $\|T_\alpha(x_0)\|, \|T_\alpha(x_0 + x)\| \leq n_0$. Thus $\|T_\alpha(x)\| \leq 2n_0$, whence $\|T_\alpha\| \leq 2n_0/r$. □

**Example 12.15** Why does $X$ have to be Banach? Suppose instead we have $X = \oplus_{n=1}^\infty \mathbb{F}$, $Y = \mathbb{F}$. Define $T_n : X \to Y$ by $T_n(e_n) = n$ and $T_n(e_i) = 1$ for all $i \neq n$. Then all $T_n$ are continuous, for each $x \in X$, $\{\|T_n(x)\| : n\}$ is bounded, but $\{\|T_n\| : n\}$ is not bounded.

**Corollary 12.16** *A subset $S$ of a normed space $X$ is bounded if and only if for each $f \in X^*$, $\{|f(s)| : s \in S\}$ is bounded (in other words, $\{f(s) : s \in S\}$ is bounded).*

*Proof.* Suppose that $S$ is bounded. Then there exists $r > 0$ such that $S \subseteq B(0, r)$. Then for all $f \in X^*$, $|f(s)| \leq \|f\| \|s\|$, so that $\{|f(s)| : s \in S\}$ is bounded.

Now assume that $\{|f(s)| : s \in S\}$ is bounded for all $f \in X^*$.

**New trick:** Not every vector space $X$ has a natural linear map $X \to X^*$, but they ALL have the natural map

$$\varphi : X \to X^{**} = (X^*)^* \text{ (“double dual”)}$$

given by $\varphi(x)(f) = f(x)$ (for $f \in X^*$). Is this even a function, i.e., is $\varphi(x) \in (X^*)^*$ for all $x$? Certainly $\varphi(x)(f + rg) = (f + rg)(x) = f(x) + rg(x) = \varphi(x)(f) + r\varphi(x)(g)$, so that $\varphi(x)$ is at least a linear functional on $X^*$. From $\|\varphi(x)\| = \sup\{|\varphi(x)(f)| : f \in X^*, \|f\| = 1\} = \sup\{|f(x)| : f \in X^*, \|f\| = 1\} \leq \|x\|$ we deduce that $\varphi(x)$ is continuous with norm at most $\|x\|$. By a consequence of the Hahn–Banach Theorem, more specifically, by Corollary 11.11, for any non-zero $x$ in $X$ (and setting $M = 0$ for the corollary), there exists $f \in X^*$ with $\|f\| = 1$ and $f(x) = \|x\|$, so that we get that $\|x\| \leq \sup\{|f(x)| : f \in X^*, \|f\| = 1\}$ $= \sup\{|\varphi(x)(f)| : f \in X^*, \|f\| = 1\} \leq \|\varphi(x)\|$, so that $\|\varphi(x)\| = \|x\|$. Thus $\varphi$ is a function from $X$ to $X^{**}$. It is also a linear function: $\varphi(x + cy)(f) = f(x + cy) = f(x) + cf(y) = \varphi(x)(f) + c\varphi(y)(f) = (\varphi(x) + c\varphi(y))(f)$, and it is injective, for if $\varphi(x) = 0$, then for all $f \in X^*$, $f(x) = 0$, so that by a consequence of the Hahn-Banach Theorem (Corollary 11.9), $x = 0$.

The assumption on $\{|f(s)| : s \in S\}$ now says that $\{|\varphi(s)(f)| : s \in S\}$ is bounded. Since $X^*$ is complete, the Uniform Boundedness Principle (Theorem 12.14) applies to give that $\{\|\varphi(s)\| : s \in S\}$ is bounded, whence by the previous paragraph, $\{\|s\| : s \in S\}$ is bounded. So $S$ is bounded. □

The last corollary has an immediate corollary: Let $X$ be a normed vector space. A subset $S$ of $X^*$ is bounded if and only if for each $f \in X^{**}$, $\{|f(s)| : s \in S\}$ is bounded. However, there is a stronger companion corollary, and it goes as follows:

**Corollary 12.17** *Let $X$ be a Banach space. Then a subset $S$ of $X^*$ is bounded if and only if for each $x \in X$, $\{|f(x)| : f \in S\}$ is bounded.*

The reason that this companion corollary is stronger is that in general $X$ is a proper subspace of $X^{**}$, However, we need to assume that $X$ is not just normed, but also complete. You provide the companion proof, and think why completeness is necessary (see Homework 8).

**Corollary 12.18** *Let $X$ be a Banach space and $Y$ a normed linear space. A subset $S \subseteq B(X, Y)$ is bounded if and only if for all $x \in X$ and all $g \in Y^*$, $\{g \circ T(x) : T \in S\}$ is bounded in $\mathbb{F}$.*

*Proof.* Suppose that $S$ is bounded. Then for all $x \in X$ and all $g \in Y^*$, $|g \circ T(x)| \leq \|g\| \|T\| \|x\|$ is bounded as $T$ varies over $S$.

Now suppose that for all $x \in X$ and all $g \in Y^*$, $\{g \circ T(x) : T \in S\}$ is bounded in $\mathbb{F}$. Then for any $x \in X$, $\{T(x) : T \in S\}$ is a subset of $Y$, and for all $g \in Y^*$, $\{g(T(x)) : T \in S\}$ is bounded, whence by Corollary 12.16, $\{T(x) : T \in S\}$ is bounded in $Y$. But then since $X$ is complete, by Theorem 12.14 (the Uniform Boundedness Principle), $S = \{T : T \in S\}$ is bounded. □

**Corollary 12.19** (The Banach–Steinhaus Theorem) *Let $X$ be a Banach space and $Y$ a normed vector space, and let $\{T_n\}$ be a sequence in $B(X, Y)$ such that for every $x \in X$ there exists $y \in Y$ such that $T_n(x)$ converges to $y$ in the norm. Then there exists $T \in B(X, Y)$ such that $T_n x \to Tx$ for all $x \in X$, and $\{T_n : n\}$ is bounded.*

*Proof.* For all $x \in X$, define a function $T : X \to Y$ such that $T(x) = \lim_n T_n(x)$. This is clearly a linear function. By the Uniform Boundedness Principle (Theorem 12.14), $\{\|T_n\| : n\}$ is bounded, say by $B$. Thus for all $x \in X$, $\|T(x)\| \leq \|T(x) - T_n(x)\| + \|T_n(x)\|$, and for $n$ sufficiently large (depending on $x$), this is at most $(1 + B) \|x\|$, so that $T$ is continuous. □

## 13 Hilbert spaces

Definition of inner product on a vector space (over $\mathbb{R}$, $\mathbb{C}$).
Every inner product defines a norm.
**Cauchy-Bunyakowsky-Schwarz inequality**:

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| .$$

A consequence is the **parallelogram law**: for all $f, g$ in an inner product space,

$$2\|f\|^2 + 2\|g\|^2 = \|f+g\|^2 + \|f-g\|^2 .$$

**Definition 13.1** *Elements $f$ and $g$ in an inner product space are* **orthogonal** *or* **perpendicular** *if $\langle f, g\rangle = 0$. Two subsets $A$ and $B$ are* **orthogonal** *or* **perpendicular** *if for all $f \in A$ and all $g \in B$, $\langle f, g\rangle = 0$. For any subset $A$, we define* **the orthogonal complement of** *$A$, or* **A "perp"**, *to be the set all elements of $H$ that are perpendicular to $A$. This set will be denoted $A^\perp$. A subset $C$ is* **orthonormal** *if for all $c \in C$, $\|c\| = 1$ and if for all distinct $c, d \in C$, $\langle c, d\rangle = 0$.*

**Definition 13.2** *A vector space is a* **Hilbert space** *if it is an inner product space that it complete in the norm induced by the inner product.*

We know that $\mathbb{F}^n$ is a Hilbert space with the inner product $\langle x, y\rangle = \sum_{i=1}^n x_i \overline{y}_i$.

More generally, the most representative example of a Hilbert space is probably $\mathbb{L}^2(X, \Sigma, \mu)$, where $(X, \Sigma, \mu)$ is a measure space. The inner product is defined as $\langle f, g\rangle = \int_X f\overline{g}\, d\mu$. First of all, if $g \in \mathbb{L}^2$, so is $\overline{g}$, and by Hölder's inequality (see page 8) or by the Cauchy-Bunyakowsky-Schwarz inequality (three paragraphs above), $f\overline{g} \in \mathbb{L}^1$, whence $\int_X f\overline{g}\, d\mu$ is well-defined. Other properties of inner products are easily verified for this $\langle\ ,\ \rangle$. We already know that $\mathbb{L}^2$ is complete in the norm.

A special and important case of $\mathbb{L}^2$ is $\ell^2 = \mathbb{L}^2(\mathbb{N}, 2^{\mathbb{N}}, \text{counting measure})$. Just like all $\ell^p$ spaces, the set $\{e_i : i \in \mathbb{N}\}$, where each $e_i$ has 1 in the $i$th spot and 0 elsewhere, is a linearly independent subset, with the property that for every $x \in \ell^p$ and every $\epsilon > 0$, there exists $e \in Span\{e_1, e_2, \ldots\}$ such that $\|x - e\| < \epsilon$. But when $p = 2$, the set $\{e_i : i \in \mathbb{N}\}$ is even orthonormal, i.e., $\langle e_i, e_j\rangle = \delta_{ij}$.

Under what conditions is $H = \mathbb{L}^p(X, \Sigma, \mu)$ also a Hilbert space? We need that all $f, g \in H$ to satisfy the parallelogram law. So for any disjoint $A, B \in \Sigma$, if $f = \chi_A$, $g = \chi_B$, the parallelogram law states that $2(\mu(A))^2 + 2(\mu(B))^2 = 2(\mu(A \cup B))^2$, i.e., that $(\mu(A))^2 + (\mu(B))^2 = (\mu(A) + \mu(B))^2$, which certainly holds if $\mu(A) = 0$ or $\mu(A) = \infty$ (and same options for $B$), but not otherwise. Thus, $\mathbb{L}^p$ is almost never a Hilbert space.

Let's do another almost $\mathbb{L}^2$ example: Let $X$ be an open subset of $\mathbb{C}$, let $\Sigma$ be the set of all Lebesgue subsets of $X$, and let $m$ be the Lebesgue measure. Then $(X, \Sigma, m)$ is a measure space, and so $\mathbb{L}^2(X, \Sigma, m)$ is a Hilbert space. The **Bergman space** $\mathbb{L}^2_a(X)$ is the subspace of all analytic functions on $X$. We will prove that $\mathbb{L}_a^2(X)$ is a closed subset of $\mathbb{L}^2(X, \Sigma, m)$, and hence a Hilbert space. So let $\{f_n\}$ be a sequence in the $\mathbb{L}_a^2(X)$ that converges to $f \in \mathbb{L}^2(X)$ in the norm. Let $a \in X$. Since $X$ is open, there exists $r > 0$ such that the closure of $B(a,r)$ is a subset of $X$. Then for all $n, m$,

$$\begin{aligned}
|(f_n - f_m)(a)| &= \left| \frac{2}{r^2} \int_{s=0}^{r} (f_n - f_m)(a) s ds \right| \\
&= \left| \frac{1}{\pi r^2} \int_{s=0}^{r} \int_{\theta=-\pi}^{\pi} (f_n - f_m)(a + se^{i\theta}) d\theta s ds \right| \\
&\qquad \text{(by the Mean Value Property for analytic functions,} \\
&\qquad \text{not covered in this course)} \\
&= \left| \frac{1}{\pi r^2} \int \int_{B(a,r)} (f_n - f_m) dm \right| \\
&= \frac{1}{\pi r^2} | \langle f_n - f_m, 1 \rangle | \text{ (inner product on } B(a,r)) \\
&\le \frac{1}{\pi r^2} \|f_n - f_m\|_2 \, \|1\|_2 \\
&\qquad \text{(by the Cauchy-Bunyakowsky-Schwartz inequality;} \\
&\qquad \text{norm on } B(a,r)) \\
&\le \frac{1}{\pi r^2} \|f_n - f_m\|_2 \sqrt{\pi r^2} \text{ (norm on } X) \\
&= \frac{1}{\sqrt{\pi} r} \|f_n - f_m\|_2 \, .
\end{aligned}$$

Let $K$ be a compact subset of $X$. Then the distance from $K$ to $\mathbb{C} \setminus X$ is positive, so that for all $a \in K$, we may take $r$ to be half of that distance. Thus on $K$, $\{f_n\}$ is Cauchy in the norm on $X$ and hence on $K$, the display above says that $\{f_n\}$ is uniformly Cauchy (also not just up to a set of measure zero). By Complex Analysis (not covered in this course), $f_n$ converges to an analytic function $g_K$ uniformly on $K$. By uniqueness of analytic functions, these $g_K$ on the various $K$ patch up to one analytic function $g$ on $X$. As in the proof of Proposition 4.7, there exists a subsequence $\{f_{n_k}\}_k$ that converges pointwise a.e. to $f$. Hence $f = g$ a.e., whence $f \in \mathbb{L}_a^2(X)$. □

## 14 Perpendicularity

**Definition 14.1** *For any vector space $X$ over $\mathbb{R}$, a subset $K$ of $X$ is* **convex** *if for all $t \in [0,1]$ and all $a, b \in K$, $ta + (1-t)b \in K$.*

Note that vector subspaces and their translates are convex.

**Theorem 14.2** *Let $H$ be a Hilbert space and let $M$ be a non-empty closed convex subset. Then for any $h \in H$ there exists a unique $x_0 \in M$ such that*

$$\|h - x_0\| = \text{dist}(h, M) = \inf\{\|h - x\| : x \in M\}.$$

*Proof.* Set $d = \text{dist}(h, M)$. By definition there exists a sequence $\{x_n\}$ in $M$ such that $\|h - x_n\| \to d$. Fix $\epsilon > 0$ and choose $N$ such that for all $n \geq N$, $|\,\|h - x_n\| - d| < \epsilon$. By the parallelogram law,

$$\|(x_n - h) - (x_m - h)\|^2 = 2\|x_n - h\|^2 + 2\|x_m - h\|^2 - \|(x_n - h) + (x_m - h)\|^2.$$

Since $M$ is convex, $(x_n + x_m)/2 \in M$, so that $\|(x_n - h) + (x_m - h)\| = 2\left\|\frac{x_n + x_m}{2} - h)\right\| \geq 2d$, whence

$$\|(x_n - h) - (x_m + h)\|^2 < 4(\epsilon + d)^2 - 4d^2 = 4\epsilon(\epsilon + 2d).$$

As we may take $\epsilon$ arbitrarily small, this says that $\{x_n - h\}$ is Cauchy in norm. Since $H$ is complete, there exists a limit $x_0$ of $\{x_n\}$ in $H$ (or $x_0 - h$ is the limit of of $\{x_n - h\}$). Since $M$ is closed, necessarily $x_0 \in M$. Then

$$d \leq \|x_0 - h\| = \|x_0 - x_n + x_n - h\| \leq \|x_0 - x_n\| + \|x_n - h\| \leq \|x_0 - x_n\| + d,$$

and in the limit as $n \to \infty$, $\|x_0 - h\| = d$, as desired.

Now suppose that $y_0 \in M$ with $\|y_0 - h\| = d$. Again by convexity, $(y_0 + x_0)/2 \in M$, so that

$$d \leq \left\|\frac{y_0 + x_0}{2} - h\right\| = \left\|\frac{y_0 - h}{2} + \frac{x_0 - h}{2}\right\| \leq \left\|\frac{y_0 - h}{2}\right\| + \left\|\frac{x_0 - h}{2}\right\| = d,$$

so that $d = \left\|\frac{y_0 + x_0}{2} - h\right\|$. But then the parallelogram law says that

$$\|y_0 - x_0\|^2 = \|(y_0 - h) - (x_0 - h)\|^2 = 2\|(y_0 - h)\|^2 + 2\|(x_0 - h)\|^2 - 2\|y_0 + x_0 - 2h\|^2 = 0,$$

so that $x_0 = y_0$ is unique. □

**Theorem 14.3** *Let $M$ be a closed linear subspace of a Hilbert space $H$. By the previous theorem, for all $h \in H$ there exists a unique $x_0 \in M$ such that $\|h - x_0\| = \text{dist}(h, M)$. Then $h - x_0$ is perpendicular to $M$.*

*Conversely, if $x_0 \in M$ has the property that $h - x_0$ is perpendicular to $M$, then $\|h - x_0\| = \text{dist}(h, M)$.*

*Proof.* First suppose that $x_0 \in H$ is for $h$ as in the previous theorem. Let $m \in M$. Then for all $t \in \mathbb{R}$, $x_0 - tm \in M$, so that $\|h - x_0 - tm\| \geq \|h - x_0\|$. In other words, $\|h - x_0\|^2 - 2tRe\langle h - x_0, m\rangle + t^2\|m\|^2 = \|h - x_0 - tm\|^2 \geq \|h - x_0\|^2$. By the previous theorem the left side is minimized at $t = 0$. By calculus or even high school algebra, such a quadratic function in $t$ is minimized at $t = 0$ if $Re\langle h - x_0, m\rangle = 0$. If $\mathbb{F} = \mathbb{R}$, we are done, otherwise similarly $Re\langle h - x_0, im\rangle = 0$, whence $\langle h - x_0, m\rangle = 0$.

Now suppose that $h - x_0$ is perpendicular to $M$. Then for all $y \in M$,

$$\begin{aligned}\|y - h\|^2 &= \|y - x_0 + x_0 - h\|^2 \\ &= \|y - x_0\|^2 + 2Re\langle y - x_0, x_0 - h\rangle + \|x_0 - h\|^2 \\ &= \|y - x_0\|^2 + \|x_0 - h\|^2 \ \text{ (since } y - x_0 \in M)\end{aligned}$$

which shows that $x_0$ is the unique element of $M$ that is closest to $h$. Thus by uniqueness from the previous theorem, this theorem is proved. □

**Theorem 14.4** *Let $M$ be a closed linear subspace of a Hilbert space $H$. Define $P : H \to H$ by $P(h)$ is the unique point in $M$ closest to $h$ (this is well-defined by Theorem 14.2). Then*

*(1) $P$ is a linear transformation.*
*(2) For all $h \in H$, $\|Ph\| \leq \|h\|$.*
*(3) $P \in B(H, H)$.*
*(4) $P^2 = P$.*
*(5) $\ker P = M^\perp$, and the range of $P$ is $M$.*

*Proof.* Let $x, y \in H$, $r \in \mathbb{F}$, and $z \in M$. Then by the previous theorem,

$$\langle x + ry - P(x) - rP(y), z\rangle = \langle x - P(x), z\rangle + r\langle y - P(y), z\rangle = 0,$$

so that by Theorem 14.3, $P(x + ry) = P(x) + rP(y)$. This proves that $P$ is linear.

For all $h \in H$, $\|Ph\|^2 = \langle Ph - h + h, Ph\rangle = \langle Ph - h, Ph\rangle + \langle h, Ph\rangle = \langle h, Ph\rangle = |\langle h, Ph\rangle| \leq \|h\| \|Ph\|$, so that if $Ph \neq 0$, we get that $\|Ph\| \leq \|h\|$. But $\|Ph\| \leq \|h\|$ holds even if $Ph = 0$, so that (2) holds.

(1) and (2) immediately imply (3).

Note that for all $h \in M$, $Ph = h$, so that for all $h \in M$, $P^2(h) = P(Ph) = Ph$ as $Ph \in M$. Thus $P^2 = P$.

The fact that for all $h \in M$, $Ph = h$, implies that the range of $P$ is $M$.

Let $h \in \ker P$, and let $m \in M$. Then $\langle h, m\rangle = \langle h - 0, m\rangle = \langle h - P(h), m\rangle = 0$ by Theorem 14.3. Thus $\ker P \subseteq M^\perp$. Now let $h \in M^\perp$. Then $\|P(h)\|^2 = \langle P(h), P(h)\rangle = \langle P(h) - h + h, P(h)\rangle = \langle P(h) - h, P(h)\rangle + \langle h, P(h)\rangle = 0$ by the previous theorem and the assumption, so that $P(h) = 0$. Thus $M^\perp \subseteq \ker P$. □

The function $P$ in the theorem above is called **the orthogonal projection of** $H$ onto $M$. Whenever $M$ is a closed linear subspace of a Hilbert space $H$, we denote by $P_M$ the corresponding orthogonal projection.

**Theorem 14.5** *If $M$ is a closed linear subspace of a Hilbert space $H$, then $I - P_M$ is an orthogonal projection onto $M^\perp$.*

*Proof.* For all $h \in H$,

$$h - (I - P_M)(h) = P_M(h) \in M$$

is perpendicular to $M^\perp$, so that by the last two theorems, $I - P_M$ is the orthogonal projection onto $M^\perp$. $\square$

**Corollary 14.6** *If $M$ is a closed linear subspace of a Hilbert space $H$, then $(M^\perp)^\perp = M$.*

*Proof.* Since $I - P_M$ is the orthogonal projection onto $M^\perp$, by Theorem 14.4 (4), $\ker(I - P_M) = (M^\perp)^\perp$. But $0 = (I - P_M)(h) = h - P_M(h)$ if and only if $h = P_M(h)$, i.e., if and only if $h \in M$, so that $M = (M^\perp)^\perp$. $\square$

**Corollary 14.7** *Let $A$ be any subset of a Hilbert space $H$. Then the closure of the linear span of $A$ equals $(A^\perp)^\perp$.*

*Proof.* Homework 8. $\square$

**Theorem 14.8** *For any closed linear subspace $M$ of a Hilbert space $H$, $H = M \oplus M^\perp$. (The inner products are the same.)*

*Proof.* If $h \in M \cap M^\perp$, then $\|h\|^2 = \langle h, h \rangle = 0$, so that $h = 0$. For any $h \in H$, $P_M(h) \in M$ and $(I - P_M)(h) \in M^\perp$, and $h = P_M(h) + (I - P_M)(h)$. If $h \in M$ and $h' \in M^\perp$, then $\|h + h'\|^2 = \|h\|^2 + 2\,\mathrm{Re}\,\langle h, h' \rangle + \|h'\|^2 = \|h\|^2 + \|h'\|^2$, which finishes the proof. $\square$

A consequence is that it is relatively easy to find orthogonal/direct-sum complements of closed linear subspaces in a Hilbert space. (It is not that easy to find direct-sum complements of closed linear subspaces in a Banach space.)

**Proposition 14.9** *Let $X$ and $Y$ be inner product spaces, and let $U \in B(X, Y)$. The following are equivalent:*

*(1) $U$ is an isomorphism (of inner product spaces).*

*(2) $U$ is surjective, and for all $x, y \in X$, $\langle U(x), U(y) \rangle = \langle x, y \rangle$.*

*(3) $U$ is surjective, and for all $x \in X$, $\|U(x)\| = \|x\|$.*

*Proof.* Clearly (1) implies (2).

(2) implies (3):

$$\begin{aligned} 2Re\,\langle U(x), U(y) \rangle &= \|U(x) + U(y)\|^2 - \|U(x)\|^2 - \|U(y)\|^2 \\ &= \|U(x + y)\|^2 - \|U(x)\|^2 - \|U(y)\|^2 \\ &= \|x + y\|^2 - \|x\|^2 - \|y\|^2 \\ &= 2Re\,\langle x, y \rangle, \end{aligned}$$

and similarly if $\mathbb{F} = \mathbb{C}$,

$$\begin{aligned} 2Im\,\langle U(x), U(y)\rangle &= -2Re(i\,\langle U(x), U(y)\rangle) \\ &= -2Re(\langle U(ix), U(y)\rangle) \\ &= -2Re\,\langle ix, y\rangle \\ &= 2Im\,\langle x, y\rangle\,. \end{aligned}$$

(3) implies (1): Clearly $U$ is injective, by assumption it is surjective, and it preserves the norm. The inverse exists and is linear, and for all $y \in Y$, $U^{-1}(y)$ has norm the same as $y$, so that $U^{-1}$ is continuous. □

## 15 The Riesz Representation Theorem

**Theorem 15.1** (The Riesz Representation Theorem) *Let $H$ be a Hilbert space. Then for any $T \in H^*$ there exists a unique $h_0 \in H$ such that for all $x \in H$,*

$$T(x) = \langle x, h_0\rangle\,.$$

*Proof.* Let $M = \ker T$. Since $T$ is continuous, $M$ is a closed linear subspace of $H$. If $M = H$, then $h_0 = 0$ is the unique element that fits the conclusion. Thus we may assume that $M \neq H$. Thus $M^\perp \neq 0$. Let $h$ be non-zero in $M^\perp$. Then $T(h) \neq 0$, and we set $h' = h/T(h)$ and $h_0 = h'/\|h'\|^2$. By linearity, $T(h_0) = 1/\|h'\|^2$. Now let $x \in H$. Then $T(x - \|h'\|^2 T(x)h_0) = T(x) - \|h'\|^2 T(x)T(h_0) = 0$, so that $x - \|h'\|^2 T(x)h_0 \in M$. Since $h_0 \in M^\perp$,

$$0 = \left\langle x - \|h'\|^2 T(x)h_0, h_0 \right\rangle = \langle x, h_0\rangle - \|h'\|^2 T(x)\langle h_0, h_0\rangle = \langle x, h_0\rangle - T(x).$$ □

**Theorem 15.2** (The Riesz–Fischer Theorem) *Let $H$ be a Hilbert space. Then $H$ is* **conjugate-linear isomorphic** *to $H^*$, or explicitly, the function*

$$\varphi : H \to H^*$$

*given by $\varphi(x)(y) = \langle y, x\rangle$ is a conjugate-linear bijective isometry.*

*Proof.* Clearly $\varphi(x) \in H^*$. For any $x, x' \in H$ and $r \in \mathbb{F}$, $\varphi(x + rx')(y) = \langle y, x + rx'\rangle = \langle y, x\rangle + \overline{r}\,\langle y, x'\rangle = \varphi(x)(y) + \overline{r}\varphi(x')(y) = (\varphi(x) + \overline{r}\varphi(x'))(y)$, so that $\varphi$ is conjugate-linear. Even though $\varphi$ is not linear, we can still talk about continuity=boundedness of conjugate linear functions! Since $|\varphi(x)(y)| = |\,\langle y, x\rangle\,| \leq \|y\|\,\|x\|$, it follows that $\|\varphi(x)\| \leq \|x\|$. Also, $\varphi(x)(x)| = \|x\|^2$ implies that $\|\varphi(x)\| \geq \|x\|$. Thus $\|\varphi(x)\| = \|x\|$ for all $x$, whence $\|\varphi\| = 1$.

In particular, it follows that $\varphi$ is injective.

Now let $f \in H^*$. By Theorem 15.1, there exists $h_0 \in H$ such that $f = \langle\ , h_0\rangle = \varphi(h_0)$. Thus $\varphi$ is surjective. □

## 16 Orthonormal bases

**Definition 16.1** *Let $H$ be a Hilbert space. A* **basis** *for $H$ is a maximal orthonormal subset of $H$.*

Certainly the standard vector space basis of $\mathbb{F}^n$ is a basis of this Hilbert space.

We will show that $\{e_i : i \in \mathbb{N}\}$ is a basis of $\ell^2$, but is NOT a vector space basis of $\ell^2$. (Silly: What is the definition of a "regular basis"?)

Orthonormal sets (of arbitrary cardinality) are linearly independent sets. Thus a straightforward application of Zorn's lemma gives the following:

**Proposition 16.2** *Any orthonormal subset (of arbitrary cardinality) of an inner product space $H$ is a subset of a vector space basis of $H$.* □

It is not clear (yet) however that every orthonormal subset of a Hilbert space $H$ is a subset of a basis of $H$.

By the standard Gram-Schmidt orthogonalization process, every countable linearly independent set can be modified into an orthonormal set that spans the same subspace.

**Proposition 16.3** *If $\{e_1, \ldots, e_n\}$ is an orthonormal set in $H$ and $M$ is the (closed) linear span of this set, then the orthogonal projection $P_M$ equals*

$$P_M(h) = \sum_{n=1}^{n} \langle h, e_n \rangle \, e_n. \qquad \square$$

*Proof.* You fill it in. □

**Theorem 16.4** (Bessel's Inequality) *If $\{e_n : n \in \mathbb{N}\}$ is an orthonormal subset of $H$, then for all $h \in H$,*

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

*Proof.* Let $h_n = h - \sum_{i=1}^{n} \langle h, e_i \rangle \, e_i$. Then $h_n$ is perpendicular to $e_i$ for all $i \leq n$, so that

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

$$\geq \sum_{i=1}^{n} |\langle h, e_i \rangle|^2.$$

Since this holds for all $n$, the theorem follows. □

**Corollary 16.5** *Let $B$ be an orthonormal subset of $H$. Then for any $h \in H$, $\langle h, b \rangle$ is non-zero for at most countably many $b \in B$.*

*Proof.* For each $n \in \mathbb{N}$, let $B_n = \{b \in B : |\langle b, h \rangle| \geq 1/n\}$. By Bessel's Inequality, the set $B_n$ is finite. But then $\cup_n B_n$ is countable, which proves the corollary. □

**Corollary 16.6** *For any orthonormal subset $B$ of a Hilbert space $H$, and any $h \in H$,*

$$\sum_{b \in B} |\langle h, b \rangle|^2 \leq \|h\|^2,$$

*and all except countably many summands are* 0. □

Is this an uncountable sum? Well, only countably many elements are non-zero, and the sum for those countably many elements converges.

But we next make this precise via **nets**.

**Definition 16.7** *A set $M$ is a* **directed set** *or a* **directed system** *if there exists a reflexive and transitive partial order $\leq$ on $M$ such that for all $a, b \in M$ there exists $c \in M$ such that $a \leq c$ and $b \leq c$.*

**Examples 16.8**

(1) $\mathbb{N}$ is a directed set under the usual $\leq$.
(2) $\mathbb{R}$ is a directed set under the usual $\leq$.
(3) $\mathbb{C}$ is a directed set under the following order: if $a, b, c, d$ are real numbers, then $a + bi \leq c + di$ if $a \leq c$ and $b \leq d$.
(4) For any set $S$, let $M$ be the set of all subsets of $S$. Then $M$ is a directed system if $\leq$ stands for set inclusion.
(5) For any set $S$, let $M$ be the set of all finite subsets of $S$. Then $M$ is a directed system if $\leq$ stands for set inclusion.
(6) (Special case of above.) Let $a, b$ be real numbers with $a \leq b$. Let $M$ be the set of all partitions of $[a, b]$ (recall calculus!). We order partitions $P \leq Q$ if $Q$ is a refinement of $P$. Then $M$ is a directed system.

**Definition 16.9** *A* **net** *is a function from a directed system to a topological space.*

Notation: Rather than writing $g(m)$ for each $m$ in the directed system, we often write $g_m$. In particular, if the directed system is $\mathbb{N}$, then a net is just a sequence.

**Definition 16.10** *We say that a net $g : (M, \leq) \to X$* **converges** *to $x \in X$ if for all neighborhoods $U$ of $x$ there exists $m_0 \in M$ such that for all $m \geq m_0$, $g(m) \in U$.*

**Examples 16.11**

(1) If $M = \mathbb{N}$, a net $g$ converges to $x$ if and only if the sequence $g$ converges to $x$.

(2) Let $M$ be the set of all partitions on $[a, b]$. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Define $U(f) : M \to \mathbb{R}$ be defined as the upper sum over partitions, let $L(f) : M \to \mathbb{R}$ be defined as the lower sum over partitions. and let $g : M \to \mathbb{R}$ satisfy that for all $P \in M$, $L(f)(P) \leq g(P) \leq U(f)(P)$. By calculus, the nets $U(f)$ and $L(f)$ converge; and if $U(f)$ and $L(f)$ converge to the same limit, then the net converges to the same limit.

Now we are ready to go back to Corollary 16.6. What do we mean by $\sum_{b \in B} a_b$?

**Definition 16.12** *Let $B$ be a set, and let $M$ be the collection of all finite subsets of $B$. Then $M$ is a directed set. Let $A : B \to X$ be a function to a normed vector space $X$. For all $F \in M$, define $A_F = \sum_{b \in F} A_b$. Thus, $A$ is a net from $M$ to a normed vector space $X$. Then by $\sum_{b \in B} A_b$ we mean the limit of the net $A$ (and this limit exists).*

**Warning:** Saying that the net $\sum_{b \in \mathbb{N}} A_b$ (with the directed system being the set of all finite subsets of $\mathbb{N}$) converges is not the same as saying that the series $\sum_{b=1}^{\infty} A_b$ converges. It is straightforward to show that the convergence of the net implies the convergence of the series. The converse fails! Let $A_b = (-1)^n/n$. By the Alternating Series Test, $\sum_{b=1}^{\infty} A_b$ converges. However, for any finite subset $m_0$, we can take sets $m = m_0 \cup \{\text{many, but finitely many, odd terms}\}$, and as we vary over such $m$, the sums $A_m$ diverge to $-\infty$. If instead we take $m = m_0 \cup \{\text{many, but finitely many, even terms}\}$, then the sums $A_m$ diverge to $\infty$. And we can take other choices of finite $m$ with yet another limit. Thus the net $\sum_{b \in \mathbb{N}} A_b$ does not converge.

**Notation 16.13** Let's agree that $\sum_{b \in \mathbb{N}} A_b$ stands for summing over the usual net with directed system being the natural order on $\mathbb{N}$, namely, $\sum_{b \in \mathbb{N}} A_b = \sum_{b=1}^{\infty} A_b$, but if we want to sum when the net is taken over the directed subset of finite subsets of $\mathbb{N}$, we will have to say so explicitly (as we did in the previous paragraph).

We repeat Corollary 16.6 with the net formulation:

**Corollary 16.14** *For any orthonormal subset $B$ of a Hilbert space $H$, and any $h \in H$, the net $\sum_{b \in B} |\langle h, b \rangle|^2$ over the directed set of all finite subsets of $B$ converges to a real number that is at most $\|h\|^2$.*

*Proof.* By Corollary 16.5 there exists a countable subset $B'$ of $B$ such that for all $b \in B \setminus B'$, $\langle h, b \rangle = 0$. Let $B' = \{e_1, e_2, \ldots\}$ be an enumeration of $B'$. By Bessel's inequality, $\sum_{i=1}^{\infty} |\langle h, e_i \rangle|^2$ converges to a real number that is at most $\|h\|^2$. Thus the countable sum $\sum_{i=1}^{\infty} |\langle h, e_i \rangle|^2$ of non-negative real numbers converges to some real number. For every

$\epsilon > 0$ there exists $N$ such that for all $n \geq N$, $\sum_{i=N+1}^{\infty} |\langle h, e_i \rangle|^2 < \epsilon$. Now let $m_0 = \{e_1, \ldots, e_N\}$. Then for any finite subset $m$ of $B$ that contains $m_0$,

$$\begin{aligned}\left|\sum_{b\in m} |\langle h, b\rangle|^2 - \sum_{i=1}^{\infty} |\langle h, e_i\rangle|^2\right| &\leq \left|\sum_{b\in m\setminus B'} |\langle h, b\rangle|^2 - \sum_{b\in B'\setminus m} |\langle h, b\rangle|^2\right| \\ &= \left|-\sum_{b\in B'\setminus m} |\langle h, b\rangle|^2\right| \\ &\leq \sum_{i=N+1}^{\infty} |\langle h, e_i\rangle|^2 \\ &< \epsilon. \qquad \square\end{aligned}$$

**Lemma 16.15** *Let $B$ be an orthonormal basis of a Hilbert space $H$. Then for all $h \in H$,*

$$\sum_{b\in B} \langle h, b\rangle\, b$$

*converges in $H$ (as a net with the directed set being the collection of finite subsets of $B$).*

*Proof.* This is a continuation of the previous proof – see set-up there.

For every finite subset $m$ of $B$, define $A_m = \sum_{b\in m} \langle h, b\rangle\, b$. In particular, for the given $\epsilon > 0$, if $m, m'$ contain $m_0$, then

$$|A_m - A_{m'}|^2 = \left|A_{m\setminus m'} - A_{m'\setminus m}\right|^2 \leq \left(\left|A_{m\setminus m'}\right| + \left|A_{m'\setminus m}\right|\right)^2 \leq \sum_{i=N+1}^{\infty} |\langle h, e_i\rangle\, e_i|^2,$$

which is less than $\epsilon$ by the previous corollary (or its proof). Thus **by definition** (which I never wrote out explicitly), $\sum_{b\in B} \langle h, b\rangle\, b$ is Cauchy in the net. Similarly, the sequence $\{\sum_{i=1}^{n} \langle h, e_i\rangle\, e_i\}_n$ is a Cauchy sequence in $H$, and as $H$ is complete, this latter sequence converges in $H$, and clearly it converges to $\sum_{i=1}^{\infty} \langle h, e_i\rangle\, e_i$. Now back to the Cauchy net $\sum_{b\in B} \langle h, b\rangle\, b$. It converges to $\sum_{i=1}^{\infty} \langle h, e_i\rangle\, e_i$ because for all finite subsets $m$ of $B$ that contain $m_0$,

$$\left|A_m - \sum_{i=1}^{\infty} \langle h, e_i\rangle\, e_i\right|^2 \leq \left|\sum_{\text{some } i=N+1}^{\infty} \langle h, e_i\rangle\, e_i\right|^2 = \sum_{\text{some } i=N+1}^{\infty} |\langle h, e_i\rangle|^2 < \epsilon. \qquad \square$$

**Theorem 16.16** *Let $B$ be an orthonormal subset of a Hilbert space $H$. Then the following are equivalent:*

*(1) $B$ is a (Hilbert space) basis of $H$.*

*(2) If $h \in H$ is perpendicular to $B$, then $h = 0$.*

*(3) The closed linear span of $B$ is $H$.*

*(4) For all $h \in H$, $h = \sum_{b \in B} \langle h, b \rangle b$.*

*(5) For all $g$ and $h$ in $H$, $\langle g, h \rangle = \sum_{b \in B} \langle g, b \rangle \langle b, h \rangle$.*

*(6)* (Parseval's Identity) *For all $h \in H$, $\|h\|^2 = \sum_{b \in B} | \langle h, b \rangle |^2$.*

*(All sums are countable in this theorem, but the proofs below are based more generally on nets.)*

*Proof.* (1) $\Rightarrow$ (2): If $h \neq 0$, then $B \cup \{h/\|h\|\}$ is an orthonormal set properly containing $B$, which is a contradiction.

(2) $\Leftrightarrow$ (3): By Theorem 14.8, the closed linear span of $B$ is $H$ if and only if the orthogonal complement of the closed linear span of $B$ is 0.

(2), (3) $\Rightarrow$ (4): By Lemma 16.15, $f = h - \sum_{b \in B} \langle h, b \rangle b \in H$, and one can show (with nets) that $f$ is perpendicular to all elements of $B$. (Similar proof in (4) $\Rightarrow$ (5).) Thus $f = 0$, which proves (4).

(4) $\Rightarrow$ (5): This is sort of obvious, except possibly for why this converges in the net. So again assume the set-up of the proof of Corollary 16.14. Let $m$ be a finite subset of $B$ containing $m_0$. Then

$$\begin{aligned}
\left| \langle g, h \rangle - \sum_{b \in m} \langle g, b \rangle \langle b, h \rangle \right| &= \left| \langle g, h \rangle - \sum_{b \in m} \left\langle g, \overline{\langle b, h \rangle} b \right\rangle \right| \\
&= \left| \langle g, h \rangle - \left\langle g, \sum_{b \in m} \langle h, b \rangle b \right\rangle \right| \\
&= \left| \left\langle g, h - \sum_{b \in m} \langle h, b \rangle b \right\rangle \right| \\
&\leq \|g\| \left\| h - \sum_{b \in m} \langle h, b \rangle b \right\|,
\end{aligned}$$

which can be made arbitrarily small.

(5) $\Rightarrow$ (6): immediate.

(6) $\Rightarrow$ (1): Let $h \in H$ be such that $B \cup \{h\}$ is orthonormal. Then $\|h\|^2 = \sum_{b \in B} | \langle h, b \rangle |^2 = 0$, which is a contradiction. $\square$

## 17 Section I.5 in the book

I am skipping it because I did much of this in Math 321 in the fall. Namely, the section proves that any two separable Hilbert spaces are isomorphic to $\ell^2$, and it proves that $\{e^{int}/(2\pi)^{1/2} : n \in \mathbb{Z}\}$ is a Hilbert basis of $\mathbb{L}^2_{\mathbb{C}}([0, 2\pi])$. It covers Fourier coefficients and Fourier series. It's a great section to know.

## 18 Adjoints of linear transformations

**Theorem 18.1** *Let $X$ be a Hilbert space, $Y$ an inner product space, and $T \in B(X, Y)$. Then there exists a unique $T^* \in B(Y, X)$ such that for all $x \in X$ and all $y \in Y$,*

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle .$$

*Proof.* Note that for all $y \in Y$, $x \mapsto \langle T(x), y\rangle$ is an element of $X^*$. Thus by the Riesz Representation Theorem (Theorem 15.1), there exists a unique $h_y \in X$ such that for all $x \in X$,

$$\langle T(x), y\rangle = \langle x, h_y\rangle .$$

Note that for all $y, y' \in Y$, $r \in \mathbb{F}$, and all $x \in X$,

$$\begin{aligned}
\langle T(x), y + ry'\rangle &= \langle T(x), y\rangle + \overline{r} \langle T(x), y'\rangle \\
&= \langle x, h_y\rangle + \overline{r} \langle x, h_{y'}\rangle \\
&= \langle x, h_y + rh_{y'}\rangle ,
\end{aligned}$$

so that by uniqueness, $h_y + rh_{y'} = h_{y+ry'}$. It follows that $y \mapsto h_y$ is a linear function from $Y$ to $X$. Since

$$\|h_y\|^2 = \langle h_y, h_y\rangle = \langle T(h_y), y\rangle = |\langle T(h_y), y\rangle| \le \|T(h_y)\| \|y\| \le \|T\| \|h_y\| \|y\| ,$$

we have that $\|h_y\| \le \|T\| \|y\|$. It follows that $y \mapsto h_y$ is in $B(Y, X)$. We label this function $T^*$. □

The proof actually gives more: for each $y$ there exists a unique $T^*(y)$ that works, and $T^*$ is linear and continuous. (Not just that there are many functions of the desired form, and only one of them is linear.)

**Definition 18.2** *Let $X$ and $Y$ be inner product spaces, and let $T : X \to Y$ and $T^* : Y \to X$ be linear functions such that for all $x \in X$ and all $y \in Y$, $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$, Then $T^*$ is called* **the adjoint of** $T$.

We just proved that adjoints of continuous linear functions from one Hilbert space to an inner product space exist.

Here is another look at adjoints, via the Riesz Representation Theorem (Theorem 15.1 and via the Riesz–Fischer Theorem (Theorem 15.2), if $X$ and $Y$ are both Hilbert spaces:

$$\begin{array}{ccc} Y & \xleftarrow{T} & X \\ y \mapsto \langle\ , y\rangle \downarrow & & \downarrow x \mapsto \langle\ , x\rangle \\ Y^* & \xrightarrow{\langle\ ,y\rangle \mapsto \langle\ , T^*(y)\rangle} & X^* \end{array}$$

This diagram commutes.

**Proposition 18.3** *Let $X, Y, Z$ be Hilbert spaces.*

*(1) For $T, S \in B(X, Y)$ and $r \in \mathbb{F}$, $(T + rS)^* = T^* + \overline{r}S^*$.*

*(2) For $T \in B(X, Y)$ and $S \in B(Y, Z)$, $(S \circ T)^* = T^* \circ S^*$.*

*(3) For $T \in B(X, Y)$, $(T^*)^* = T$.*

*(4) For an invertible $T \in B(X, Y)$, $T^*$ is also invertible, and $(T^*)^{-1} = (T^{-1})^*$.*

*Proof.* (1) For all $x \in X$ and $y \in Y$,

$$\begin{aligned} \langle (T + rS)(x), y\rangle &= \langle T(x), y\rangle + r\,\langle S(x), y\rangle \\ &= \langle x, T^*(y)\rangle + r\,\langle x, S^*(y)\rangle \\ &= \langle x, T^*(y)\rangle + \langle x, \overline{r}S^*(y)\rangle \\ &= \langle x, T^*(y) + \overline{r}S^*(y)\rangle \\ &= \langle x, (T^* + \overline{r}S^*)(y)\rangle, \end{aligned}$$

so that by uniqueness, $(T + rS)^* = T^* + \overline{r}S^*$.

(2) For all $x \in X$, $y \in Y$, and $z \in Z$,

$$\langle (S \circ T)(x), z\rangle = \langle S(T(x)), z\rangle = \langle T(x), S^*(z)\rangle = \langle x, T^*(S^*(z))\rangle = \langle x, (T^* \circ S^*)(z)\rangle,$$

and again uniqueness proves (2).

(3) also follows from uniqueness and the fact that for all $x \in X$, $y \in Y$,

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle = \overline{\langle T^*(y), x\rangle} = \overline{\langle y, (T^*)^*(x)\rangle} = \langle (T^*)^*(x), y\rangle.$$

Thus for all $y \in Y$, $\langle T(x) - T^{**}(x), y\rangle = 0$, so that $T(x) - T^{**}(x) = 0$.

(4) By the Inverse Function Theorem (Theorem 12.9), $T^{-1}$ is also continuous, so its adjoint exists as well. Then $I = I^* = (T^{-1} \circ T)^* = T^* \circ (T^{-1})^*$, and $I = (T^{-1})^* \circ T^*$, so that $T^*$ is the inverse of $(T^{-1})^*$. □

**Examples 18.4**

(1) The adjoint of the identity is the identity.

(2) If $T : \ell^2 \to \ell^2$ is given by the **unilateral shift** $T(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots)$, then $T \in B(\ell^2, \ell^2)$. Since (with $e_n$ the standard element of $\ell^2$)

$$\langle e_n, T^*(e_m)\rangle = \langle T(e_n), e_m\rangle = \langle e_{n+1}, e_m\rangle = \begin{cases} 1, & \text{if } n+1 = m; \\ 0, & \text{otherwise.} \end{cases}$$

by uniqueness necessarily $T^*(e_m) = e_{m-1}$, i.e., $T^*$ is the **backward shift** (and it is a continuous linear function).

(3) If $T : \ell^2 \to \ell^2$ is given by $T(x_1, x_2, x_3, \ldots) = (x_2, x_4, x_6, \ldots)$, then $T \in B(\ell^2, \ell^2)$. Since

$$\begin{aligned}\langle (x_2, x_4, x_6, \ldots), (y_1, y_2, y_3, \ldots)\rangle &= \langle T(x_1, x_2, x_3, \ldots), (y_1, y_2, y_3, \ldots)\rangle \\ &= \langle (x_1, x_2, x_3, \ldots), T^*(y_1, y_2, y_3, \ldots)\rangle,\end{aligned}$$

by uniqueness necessarily $T^*(y_1, y_2, y_3, \ldots) = (0, y_1, 0, y_2, 0, y_3, \ldots)$.

(4) Let $T \in (\ell^2)^*$ be given by $T(x_1, x_2, x_3, \ldots) = x_1$. From

$$\langle x_1, y\rangle = \langle T(x_1, x_2, x_3, \ldots), y\rangle = \langle (x_1, x_2, x_3, \ldots), T^*(y)\rangle$$

we deduce that $T^*(y) = (y, 0, 0, 0, \ldots)$.

## 19 Self-adjoint operators

The goal is to build to a spectral decomposition of (some) operators.

**Definition 19.1** *Let $X$ be an inner product space. An element $T \in B(X)$ is* **self-adjoint** *if $T^* = T$, and it is* **skew-adjoint** *if $T^* = -T$.*

If $X$ is a Hilbert space, and $T \in B(X)$, $T - T^*$ is skew-adjoint, and the following are self-adjoint: $T \circ T^*$, $T^* \circ T$, $T + T^*$.

**Proposition 19.2** *If $X$ is a Hilbert space, every $T \in B(X)$ can be written uniquely as $T = T_1 + T_2$, where $T_1, T_2 \in B(X)$ and $T_1$ is self-adjoint and $T_2$ is skew-adjoint. If $\mathbb{F} = \mathbb{C}$, then $T$ can be written uniquely as $T = S_1 + iS_2$, where both $S_1$ and $S_2$ are self-adjoint.*

*Proof.* Let $T_1 = \frac{1}{2}(T + T^*)$ and $T_2 = \frac{1}{2}(T - T^*)$. This proves the existence of the decomposition $T = T_1 + T_2$. Suppose that $T = T_1' + T_2'$ for some self-adjoint $T_1'$ and skew-adjoint $T_2'$ in $B(X)$. Then $T_1 - T_1'$ is self-adjoint and equals $T_2' - T_2$, so that $T_2' - T_2 = (T_2' - T_2)^* = (T_2')^* - T_2^* = -T_2' + T_2$, whence $T_1 - T_1' = T_2' - T_2 = 0$. This proves uniqueness.

If $\mathbb{F} = \mathbb{C}$, set $S_1 = T_1$ and $S_2 = -iT_2$. Certanly $S_1$ is self-adjoint, and $S_2^* = (-iT_2)^* = \overline{-i}T_2^* = i(-T_2) = -iT_2 = S_2$, so that $S_2$ is self-adjoint as well. You verify uniqueness. □

**Proposition 19.3** *An orthogonal projection onto a closed subspace of a Hilbert space is self-adjoint.*

*Proof.* Let $X$ be a Hilbert space and $M$ a closed subspace. Let $x, y \in X$. Then

$$\begin{aligned}
\langle P_M(x), y\rangle &= \langle P_M(x), (y - P_M(y)) + P_M(y)\rangle \\
&= \langle P_M(x), y - P_M(y)\rangle + \langle P_M(x), P_M(y)\rangle \\
&= \langle P_M(x), P_M(y)\rangle \\
&= \langle P_M(x), P_M(y)\rangle + \langle x - P_M(x), P_M(y)\rangle \\
&= \langle P_M(x) + x - P_M(x), P_M(y)\rangle \\
&= \langle x, P_M(y)\rangle ,
\end{aligned}$$

so that $P_M^* = P_M$. □

**Theorem 19.4** *Let $X$ be a Hilbert space and $T \in B(X)$. Then $T$ is an orthogonal projection (onto some closed subspace) if and only if $T = T^* = T^2$.*

*Proof.* One direction has been proved. So we assume that $T = T^* = T^2$. Let $M$ be the range of $T$. If $x \in \overline{M}$, then there exists a sequence $\{x_n\}_n$ in $M$ that converges to $x$ in the norm. Then by continuity of $T$ and since $T^2 = T$, $x_n = T(x_n)$ converges to $x = T(x)$, so that $M$ is closed. For $x \in H$ and any $y \in M$,

$$\langle x - T(x), y\rangle = \langle x, y\rangle - \langle T(x), y\rangle = \langle x, y\rangle - \langle x, T^*(y)\rangle = \langle x, y\rangle - \langle x, T(y)\rangle = \langle x, y\rangle - \langle x, y\rangle = 0,$$

so that $x - T(x)$ is perpendicular to $M$, so that by Theorem 14.3, $T = P_M$. □

**Proposition 19.5** *Let $H$ be a complex Hilbert space. An element $T \in B(H)$ is self-adjoint if and only if for all $x \in H$, $\langle Tx, x\rangle$ is real.*

*Proof.* If $T = T^*$, then for all $x \in H$, $\langle T(x), x\rangle = \langle x, T^*(x)\rangle = \langle x, T(x)\rangle = \overline{\langle T(x), x\rangle}$, so that $\langle T(x), x\rangle$ is real.

For all $x, y \in H$,

$$\begin{aligned}
4\langle T(x), y\rangle &= \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle + i\langle T(x+iy), x+iy\rangle - i\langle T(x-iy), x-iy\rangle , \\
4\langle x, T(y)\rangle &= \langle x+y, T(x+y)\rangle - \langle x-y, T(x-y)\rangle + i\langle x+iy, T(x+iy)\rangle - i\langle x-iy, T(x-iy)\rangle .
\end{aligned}$$

If all the inner products on the right sides are real, then the two right sides are identical, proving that $\langle T(x), y\rangle = \langle x, T(y)\rangle$, so that by uniqueness of adjoints, $T = T^*$, so $T$ is self-adjoint. □

**Definition 19.6** *Let $T, S$ be a self-adjoint continuous operators on a Hilbert space $H$. We say that $T \leq S$ if for all $x \in H$, $\langle T(x), x\rangle \leq \langle S(x), x\rangle$.*

Clearly, $T \leq \|T\| I$. Thus there exists a least real number $C \in \mathbb{R}$ such that $T \leq CI$. This $C$ is called the **least upper bound** for $T$. Similarly, since $-\|T\| I \leq T$, there exists a greatest real number $c \in \mathbb{R}$ such that $cI \leq T$. This $C$ is called the **greatest lower bound** for $T$.

The least upper bound for $T$ can be strictly smaller than $\|T\|$. Namely, if $T = -I$, then $\|T\| = 1$ and the least upper bound is $-1$.

**Proposition 19.7** *Let $T \in B(X)$, where $X$ is a Hilbert space. Then* $\sup\{|\langle T(x), x\rangle| : \|x\| = 1\} = \|T\|$.

*Proof.* Certainly the least upper bound $M$ is at most $\|T\|$. Let $x, y$ be in the space with norm 1. Then

$$\begin{aligned} 4|Re\,\langle T(x), y\rangle| &= |\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle| \\ &\leq |\langle T(x+y), x+y\rangle| + |\langle T(x-y), x-y\rangle| \\ &\leq M(|\langle x+y, x+y\rangle| + |\langle x-y, x-y\rangle|) \\ &= 2M(\|x\|^2 + \|y\|^2) \\ &= 4M. \end{aligned}$$

Thus if $\langle T(x), y\rangle = e^{i\theta}|\langle T(x), y\rangle|$, we have that

$$|\langle T(x), y\rangle| = Re\left\langle T(x), e^{i\theta}y\right\rangle \leq M,$$

and this holds for all $x, y$ of norm 1. If $T(x) \neq 0$, we get that

$$\left\langle T(x), \frac{T(x)}{\|T(x)\|}\right\rangle \leq M,$$

i.e., that $\|T(x)\| \leq M$. But the last inequality holds even if $T(x) = 0$. Since $x$ varies over all elements of norm 1, this proves that $\|T\| \leq M$. □

Let $X$ be a Hilbert space and $T \in B(X)$. Let $x_n \in X$ such that $\|x_n\| = 1$ and $|\langle T(x_n), x_n\rangle| \to \|T\|$ (such a sequence exists by the previous proposition). Since the set $\{\alpha \in \mathbb{F} : |\alpha| = \|T\|\}$ is compact, by taking subsequence we may assume that $\{\langle T(x_n), x_n\rangle\}_n$ converges to $e^{i\theta}\|T\|$ for some real $\theta$. But all $\langle T(x_n), x_n\rangle$ are real, so that $|\langle T(x_n), x_n\rangle| \to \pm\|T\|$. Then

$$0 \leq \|T(x_n) \mp \|T\| x_n\|^2 = \|T(x_n)\|^2 \mp 2\|T\|\langle T(x_n), x_n\rangle) + \|T\|^2\|x_n\|^2 \leq \|T\|^2 \mp 2\|T\|\langle T(x_n), x_n\rangle + \|T\|^2,$$

and this converges to $\|T\|^2 - 2\|T\|^2 + \|T\|^2 = 0$, which proves that $\|T\| \in \sigma_{ap}(T)$ (definition in the next section) but $\|T\|$ need not be an eigenvalue of $T$. See next section.

## 20 Spectrum

Let $X$ be a non-zero complex vector space, and let $T \in B(X)$. If $p$ is a polynomial in one variable, with $p(z) = a_0 + a_1 z + \cdots + a_n z^n$, define

$$p(T) = a_0 I + a_1 T + \cdots + a_n T^n.$$

Obviously, $p_1(T)p_2(T) = (p_1 p_2)(T)$.

**Definition 20.1** *For $T \in B(X)$,*

$$\begin{aligned} \sigma(T) &= \textbf{spectrum of } T = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not invertible}\}, \\ \sigma_p(T) &= \textbf{point-spectrum of } T = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not one-to-one}\}, \\ \rho(T) &= \textbf{resolvent set of } T = \mathbb{C} \setminus \sigma(T). \end{aligned}$$

*Also, the* **resolvent function** *is $R(\lambda, T) = (\lambda I - T)^{-1}$ (for $\lambda \in \rho(T)$). Elements of the point-spectrum are called the* **eigenvalues** *of $T$. A complex number $\lambda$ is an eigenvalue of $T$ if and only if there exists a non-zero $x \in X$ such that $T(x) = \lambda x$. Any such $x$ is called an* **eigenvector** *of $T$ corresponding to the eigenvalue $\lambda$.*

*An* **approximate eigenvector** *of $T$ corresponding to $\lambda$ is a sequence $\{x_n\}$ in $X$ such that $\|x_n\| = 1$ for all $n$ and such that $T(x_n) - \lambda x_n \to 0$ in the norm. Define*

$$\begin{aligned} \sigma_{ap}(T) &= \textbf{an approximate spectrum of } T \\ &= \{\lambda \in \mathbb{C} : \text{ there exists an approximate eigenvector corresponding to } \lambda\}. \end{aligned}$$

Certainly $\sigma_p \subseteq \sigma_{ap}$ and $\sigma_p \subseteq \sigma$. If $X$ is finite-dimensional, $\sigma_p(T) = \sigma_{ap}(T) = \sigma(T)$.

**Proposition 20.2** *Let $X$ be a Hilbert space, and let $T \in B(X)$. Then $\sigma(T^*) = \{\overline{z} : z \in \sigma(T)\}$.*

*Proof.* Write it out. □

**Theorem 20.3** *Let $X$ be a Banach space, $T \in B(X)$, and $\lambda \in \mathbb{C}$. The following are equivalent:*

*(1) $\lambda \notin \sigma_{ap}(T)$.*

*(2) There exists $c > 0$ such that for all $x \in X$, $\|(\lambda I - T)(x)\| \geq c\|x\|$.*

*(3) $\lambda I - T$ is one-to-one with closed range.*

*Proof.* (1) ⇒ (2): Suppose that for all $n \in \mathbb{N}$ there exists $x_n \in X$ such that $\|(\lambda I - T)(x_n)\| < \frac{1}{n}\|x_n\|$. Then $x_n \neq 0$ and $\|(\lambda I - T)(y_n)\| = \|(\lambda I - T)(x_n/\|x_n\|)\| < \frac{1}{n}$, so that $\lambda \in \sigma_{ap}(T)$ (with approximate eigenvector $\{x_n/\|x_n\|\}$).

(2) ⇒ (3): Let $x \in \operatorname{Ker}(\lambda I - T)$. Then $0 = \|(\lambda I - T)(x)\| \geq c\|x\|$ implies that $x = 0$. Thus $\lambda I - T$ is one-to-one. Let $x$ be in the closure of the range of $\lambda I - T$. Let

$y_n = (\lambda I - T)(y'_n)$ be such that $\|y_n - x\| < 1/n$. Then $\{y_n\}$ is a Cauchy sequence, so by assumption (2), $\{y'_n\}$ is a Cauchy sequence in $X$. Since $X$ is Banach, there exists $y \in X$ such that $y'_n \to y$. By continuity of $\lambda I - T$, $(\lambda I - T)(y) = \lim_{n\to\infty}(\lambda I - T)(y'_n) = \lim_{n\to\infty} y_n = x$. Thus the range of $\lambda I - T$ is closed.

(3) $\Rightarrow$ (1): Let $Y$ be the range of $\lambda I - T$. Then $\lambda I - T : X \to Y$ is a bijective continuous linear mapping of Banach spaces, so that by the Inverse Mapping Theorem, there exists $U \in B(Y, X)$ such that $U \circ (\lambda I - T) = I$ and $(\lambda I - T) \circ U = I$. Then for all $x \in X$ with $\|x\| = 1$, $1 = \|(U \circ (\lambda I - T))(x)\| \le \|U\| \, \|(\lambda I - T)(x)\|$, so that $\|(\lambda I - T)(x)\| \ge 1/\|U\|$, whence $\lambda \notin \sigma_{ap}(T)$. □

**Example 20.4** Let $T : \ell^2 \to \ell^2$ be the unilateral shift. Then $\sigma_p(T) = \emptyset$ (work through it). For which $\lambda$ is $\lambda I - T$ surjective? Try to solve for $x$ in $(\lambda I - T)(x) = y$:

$$(\lambda x_1, \lambda x_2 - x_1, \lambda x_3 - x_2, \ldots) = (y_1, y_2, y_3, \ldots).$$

If $\lambda = 0$, this does not have a solution for all $y$, so that $0 \in \sigma(T)$. It is an exercise (Exercise 10.1) to determine $\sigma(T)$. Observe in any case that here $\sigma_p(T)$ is a proper subset of $\sigma(T)$.

Suppose that $|\lambda| \ne 1$. Then for all $x$,

$$\|(\lambda I - T)(x)\| \ge |\, \|\lambda x\| - \|T(x)\| \,| = ||\lambda| - 1| \, \|x\|,$$

so by Theorem 20.3, $\lambda \notin \sigma_{ap}(T)$. Now let $\lambda$ have absolute value 1. Set $x_n = (1, 1/\lambda, 1/\lambda^2, \ldots, 1/\lambda^{n-1}, 0, 0, \ldots)$. Then $x_n \in \ell^2$ with $\|x_n\| = \sqrt{n}$, and $(\lambda I - T)(x_n) = (\lambda, 0, \ldots, 0, -1/\lambda^{n-1}, 0, 0, \ldots)$ has norm $(|\lambda|^2 + 1/|\lambda^{n-1}|^2)^{1/2} = 2^{1/2}$, so that $\{x_n/\|x_n\|\}$ is an approximate eigenvector for $\lambda$. Thus $\sigma_{ap}(T) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

In contrast, $\sigma_p(T^*) = \sigma(T^*) = B(0,1)$. Namely, $(\lambda I - T^*)(x_1, x_2, \ldots) = (\lambda x_1 - x_2, \lambda x_2 - x_3, \lambda x_3 - x_4, \ldots)$. If $|\lambda| < 1$, then $(1, \lambda, \lambda^2, \ldots)$ is in the kernel of $\lambda I - T^*$, so that $B(0,1) \subseteq \sigma_p(T^*)$. If $\lambda \in \mathbb{C}$ and $x \in \ker(\lambda I - T^*)$, then $x_{n+1} = \lambda x_n = \cdots = \lambda^n x_1$, so that if $x_1 \ne 0$, necessarily $(1, \lambda, \lambda^2, \ldots) \in \ell^2$, so that $\lambda \in B(0,1)$. This proves that $\sigma_p(T^*) = B(0,1)$. We already know that $\sigma_p(T^*) \subseteq \sigma(T^*) = \sigma(T)$, so see Exercise 10.1 to determine $\sigma(T^*)$.

We know that $\sigma_p(T^*) \subseteq \sigma_{ap}(T^*)$. We claim that $\sigma_{ap}(T^*) = \{\lambda \in \mathbb{C} : |\lambda| \le 1\}$. First suppose that $|\lambda| > 1$. Then for all $x \in \ell^2$,

$$\|\lambda x - T^*(x)\| \ge \|\lambda x\| - \|T^*(x)\| = |\lambda| \, \|x\| - (\|x\|^2 - |x_1|^2)^{1/2} \ge |\lambda| \, \|x\| - \|x\| = (|\lambda| - 1) \|x\|,$$

which proves that $\lambda \notin \sigma_{ap}(T)$ if $|\lambda| > 1$. Now let $\lambda \in \mathbb{C}$ have absolute value 1. Set $x_n = (1, \lambda, \lambda^2, \ldots, \lambda^{n-1}, 0, 0, \ldots)$. Then

$$\|x_n\| = \left( \sum_{i=0}^{n-1} |\lambda|^{2i} \right)^{1/2} = \sqrt{n},$$

and $\|(\lambda I - T^*)(x_n)\| = \|(0, 0, \ldots, 0, \lambda^n, 0, 0, \ldots)\| = |\lambda|^n = 1$, so that $\{x_n/\sqrt{n}\}$ is an approximate eigenvector for $\lambda$.

**Lemma 20.5** *Let $T \in B(X)$ with $\|T\| < 1$. If $X$ is a Banach space, then $I - T$ is invertible and $(I-T)^{-1} = \sum_{i=0}^{\infty} T^n$.*

*Proof.* Since $\|T^n\| \leq \|T\|^n$, the series $\sum_{i=0}^{\infty} T^n$ converges absolutely. By a homework problem, since $B(X)$ is complete, the series then converges in $B(X)$. Furthermore, the composition of this series with $I - T$ in either order gives the identity, which proves the lemma. □

**Theorem 20.6** *Let $X$ be a Banach space. Then the set of invertible elements of $B(X)$ is an open subset of $B(X)$, and the map $T \mapsto T^{-1}$ is continuous.*

*Proof.* There is nothing to prove if $X = 0$, so we assume that $X$ is non-zero. Suppose that $T$ is invertible. Then $\|T^{-1}\| > 0$. Let $S \in B(X)$ such that $\|T - S\| < 1/\|T^{-1}\|$. Then

$$S = T - (T - S) = T(I - T^{-1}(T - S)),$$

and since $\|T^{-1}(T-S)\| \leq \|T^{-1}\| \|T - S\| < 1$, by the lemma, $I - T^{-1}(T-S)$ is invertible, so that $S$ is a composition of two invertible operators, hence invertible. Thus the set of invertible operators is open. Furthermore, $S^{-1} = \sum_{n=0}^{\infty}(T^{-1}(T-S))^n T^{-1}$, and so

$$\begin{aligned}
\|S^{-1} - T^{-1}\| &= \left\| \sum_{n=1}^{\infty} (T^{-1}(T-S))^n T^{-1} \right\| \\
&\leq \sum_{n=1}^{\infty} \|T^{-1}\|^{n+1} \|T - S\|^n \\
&= \|T^{-1}\|^2 \|T - S\| \sum_{n=0}^{\infty} (\|T^{-1}\| \|T - S\|)^n \\
&= \frac{\|T^{-1}\|^2 \|T - S\|}{1 - \|T^{-1}\| \|T - S\|}.
\end{aligned}$$

As $S \to T$, this quantity gets smaller and smaller, which proves the continuity. □

**Corollary 20.7** *$\rho(T)$ is open, $\sigma(T)$ is closed.* □

**Lemma 20.8** *If $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$.*

*Proof.* $\lambda I - T = \lambda(I - T/\lambda)$, and as $\|T/\lambda\|$ is strictly smaller than 1, then $I - T/\lambda$ is invertible, so $\lambda I - T$ is invertible. □

**Corollary 20.9** *$\sigma(T) \subseteq \overline{B(0, \|T\|)}$.* □

**Theorem 20.10** *The boundary of $\sigma(T)$ is a subset of $\sigma_{ap}(T)$.*

*Proof.* Let $\lambda$ be in the boundary of $\sigma(T)$. Then there exists a sequence $\{\lambda_n\}$ in $\rho(T)$ that converges to $\lambda$. Set $S = \lambda I - T$, $S_n = \lambda_n I - T$. Then all $S_n$ are invertible, but $S$ is not as $\sigma(T)$ is closed. Thus by the proof of Theorem 20.6, $\|S - S_n\| \geq 1/\left\|S_n^{-1}\right\|$. But $\|S_n - S\| = \|(\lambda - \lambda_n)I\|$, so that $\{1/\left\|S_n^{-1}\right\|\}$ converges to 0. In other words, $\{\left\|S_n^{-1}\right\|\}$ diverges to $\infty$. Set $B_n = S_n^{-1}/\left\|S_n^{-1}\right\|$. So $\|B_n\| = 1$, and

$$\|SB_n\| = \|(S - S_n)B_n + S_nB_n\| \leq \|(S - S_n)B_n\| + \|S_nB_n\| \leq \|S - S_n\| + 1/\left\|S_n^{-1}\right\|$$

converges to 0. Since $\|B_n\| = 1$, there exists $y_n \in X$ such that $\|y_n\| = 1$ and $\|B_n(y_n)\| > 1 - 1/n$. Let $x_n = B_n(y_n)/\|B_n(y_n)\|$. As $SB_n$ converges to 0 in the norm, we have that $S(x_n)$ converges to 0 in the norm. □

**Definition 20.11** **The spectral radius** *of $T$ is*

$$r_\sigma(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

*(For now we assume that $\sigma(T) \neq \emptyset$. We'll show later that this is always true.)*

We just proved that $r_\sigma(T) \leq \|T\|$, and below we will prove that $r_\sigma(T) \leq \inf\{\|T^n\|^{1]n} : n \in \mathbb{N}\}$.

**Theorem 20.12** (Polynomial Spectral Mapping Theorem) *If $X$ is a Banach space, $T \in B(X)$, and $p$ is a polynomial in one variable of positive degree, then $\sigma(p(T)) = p(\sigma(T))$.*

*Proof.* Let $\lambda \in \mathbb{C}$. Write $p(z) - \lambda = \alpha(z - r_1)(z - r_2)\cdots(z - r_n)$ for some $\alpha, r_i \in \mathbb{C}$. (Necessarily $n = \deg p$.) Then

$$p(T) - \lambda I = \alpha(T - r_1 I)(T - r_2 I)\cdots(T - r_n I),$$

and the factors commute. Thus $\lambda \in \rho(p(T))$ if and only if $\lambda I - p(T)$ is invertible, which in turn holds if and only if $r_i I - T$ is invertible for all $i$, i.e., if and only if $r_i \notin \sigma(T)$ for all $i$. But $r_1, \ldots, r_n$ are precisely those complex numbers $z$ for which $p(z) = \lambda$. Thus if $\lambda \in \sigma(p(T))$, then some $r_i \in \sigma(T)$, so that $\lambda = p(r_i) \in p(\sigma(T))$. For the other inclusion, if $\lambda \in p(\sigma(T))$, write $\lambda = p(s)$ for some $s \in \sigma(T)$. Necessarily $s = r_i$ for some $i$, so that $r_i \in \sigma(T)$, whence by above $\lambda \in \sigma(p(T))$. This proves the theorem. □

**Corollary 20.13** *If $\lambda \in \sigma(T)$, then $\lambda^n \in \sigma(T^n)$.* □

**Theorem 20.14** *If $\lambda \in \sigma(T)$, then for all $n$, $|\lambda| \leq \|T^n\|^{1/n}$.*

*Proof.* Use the previous corollary and Corollary 20.9. □

Can we extend this to other functions? What do we mean by $f(T)$ if $f$ is not necessarily polynomial? See next section.

## 21 Holomorphic Banach space-valued functions

If $Y$ is a Banach space, $U$ an open subset of $\mathbb{C}$, and $z_0 \in U$, a function $f : U \to Y$ has **complex derivative** at $z_0$ if

$$\lim_{z \to z_0} \frac{1}{z - z_0}(f(z) - f(z_0))$$

exists (using the norm topology in $Y$). If the limit exists, we denoted it by $f'(z_0)$. We say that $f$ is **(complex) analytic** or **holomorphic** on $U$ if $f'(z)$ exists for all $z \in U$.

**Facts 21.1** Some of the facts below are straightforward, and some are only listed as a background, with many details to be checked.

(1) (Easy.) If $f$ is analytic, then $f$ is continuous.

(2) (**Weak integrals of Banach-space valued functions**) Let $\mu$ be a real- or complex-valued measure on a set $S$ and let $f : S \to Y$. We say that $\int f d\,\mu = y$ **weakly** if for all $F \in Y^*$, $F \circ f$ is integrable in the standard sense and $\int (F \circ f) d\,\mu = F(y)$.

(3) (Easy.) If $y$ as above exists, it is unique.

(4) Proposition: If $K$ is a compact metric space, $\mu$ is a (finite) complex measure on $K$, and $f : K \to Y$ is a norm-continuous function, then the weak integral $\int f d\,\mu$ exists. (The idea of the proof is for each $\epsilon > 0$ to cover $K$ with with balls of radius $\delta$ ($\delta$ such that all $x$ within $\delta$ of each other map via $f$ to within $\epsilon$ of each other), and to partition $K$ into pairwise disjoint measurable sets with diameters at most $\delta$. Approximate the integral via the partitions, and show that you do have an approximation.)

(5) (Straightforward.) If the weak integral exists and if $\|f(\ )\| : K \to \mathbb{R}$ is measurable, then

$$\left\| \int f d\,\mu \right\| \leq \int \|f(s)\|\, d\,\mu(s).$$

(6) (Goursat's lemma) If $f$ is analytic in a region containing a closed rectangular region $R$, then $\int_{\partial R} f = 0$.

(7) (Cauchy Integral Theorem) If $f$ is analytic in an open set $U$ and $C$ is a cycle in $U$ (= "sum" of finitely many oriented pairwise non-intersecting closed curves) such that the winding number $W(C, z) = 0$ for all $z \notin U$, then $\int_C f = 0$.

(8) (Cauchy Integral Formula) If $f$ is analytic in an open set $U$, $z_0 \in U$, and $C$ is a cycle in $U$ such that the winding number $W(C, z_0) = 1$ and $W(C, z) = 0$ for all $z \notin U$, then

$$f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0} dz.$$

(9) (Cauchy's inequalities) If $f$ is analytic in $\{z : |z - z_0| < r\}$ and $\|f(z)\| \leq M$ for all $z$ in this set, then

$$\left\| \frac{f^{(n)}(z_0)}{n!} \right\| \leq \frac{M}{r^n}.$$

(10) If $f$ is analytic in $\{z : |z - z_0| < r\}$, then

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

in this region.

(11) (Liouville's Theorem) If $f$ is analytic on $\mathbb{C}$ and $\{\|f(z)\| : z \in \mathbb{C}\}$ is bounded, then $f$ is constant.

(12) (Follows immediately from the proof of Theorem 20.6.) (Resolvent identity) If $z, w \in \rho(T)$, then

$$(zI-T)^{-1}-(wI-T)^{-1} = (w-z)(zI-T)^{-1}(wI-T)^{-1} = (w-z)(wI-T)^{-1}(zI-T)^{-1}.$$

(13) The resolvent function $R(\lambda, T) = (\lambda I - T)^{-1}$ is an analytic function on $\rho(T)$.

**Corollary 21.2** $\sigma(T) \neq \emptyset$.

*Proof.* If $\sigma(T) = \emptyset$, then the resolvent function of $T$ is analytic on all of $\mathbb{C}$. But $(\lambda I - T)^{-1} = \sum_{n=1}^{\infty} \lambda^{-n} T^{n-1}$ for $|\lambda| > \|T\|$. But then $\lim_{|\lambda| \to \infty} (\lambda I - T)^{-1} = 0$. By Liouville's Theorem, $(\lambda I - T)^{-1}$ is a constant function, so necessarily the zero function, which is a contradiction. □

Now we go back to the end of the previous section, and in particular to the Polynomial Spectral Theorem. We take an analytic function $f : U \to \mathbb{C}$, where $U$ is a neighborhood of $\sigma(T)$, and we choose a cycle $C$ in $U \setminus \sigma(T)$ such that the winding number of $C$ around $z$ equals

$$W(C, z) = \begin{cases} 1, & \text{if } z \in \sigma(T); \\ 0, & \text{if } z \notin U. \end{cases}$$

Then define

$$f(T) = \frac{1}{2\pi i} \int_C f(z)(zI - T)^{-1} dz.$$

Existence of such a cycle: note that $\sigma(T)$ is a closed and bounded subset of $\mathbb{C}$; learn complex analysis or even topology. If $C'$ is another cycle, then $C - C'$ is another cycle in $U \setminus \sigma(T)$, with winding numbers zero around $z \in \sigma(T)$ and around $z \notin U$. Since $f(z)(zI - T)^{-1}$ is **analytic** on $U \setminus \sigma(T)$, by **Cauchy's integral formula** (study complex analysis), the integral over $C - C'$ is 0, which verifies that $f(T)$ is well-defined.

**Facts 21.3**

(1) $(f+g)(T) = f(T) + g(T)$.
(2) $(\lambda f)(T) = \lambda f(T)$.
(3) (This requires some integral work, Cauchy Integral Formula, etc.!) $(fg)(T) = f(T)g(T)$.
(4) $f(T)g(T) = g(T)f(T)$.
(5) If $f(z) = z^n$, then $f(T) = T^n$. By above it suffices to prove this for $n = 0$, $n = 1$. (This also requires some work!)
(6) If $f(z) = e^z$, we get $f(T) = e^T$. If $T$ and $S$ commute, then $e^T e^S = e^{T+S}$.

**Lemma 21.4** *Let $X$ be a Banach space and $T \in B(X)$. If $f$ is analytic on a neighborhood of $\sigma(T)$ and if $f(z) \neq 0$ for all $z \in \sigma(T)$, then $f(T)$ is invertible and $(f(T))^{-1} = (1/f)(T)$.*

*Proof.* $(1/f)(T) \circ f(T) = ((1/f) \cdot (f))(T) = I$ by the case $n = 0$ in the last fact above. □

**Theorem 21.5** (Spectral Mapping Theorem) *Let $X$ be a Banach space, $T \in B(X)$, and $f$ analytic on a neighborhood of $\sigma(T)$. Then $\sigma(f(T)) = f(\sigma(T))$.*

*Proof.* Suppose that $\lambda \notin f(\sigma(T))$. Then $\lambda \neq f(z)$ for all $z \in \sigma(T)$. Then by Lemma 21.4, $(\lambda - f)(T) = \lambda I - f(T)$ is invertible. Thus $\lambda \notin \sigma(f(T))$.

Now suppose that $\lambda \in f(\sigma(T))$. So there exists $z_0 \in \sigma(T)$ such that $f(z_0) = \lambda$. Let $g(z) = (f(z) - \lambda)/(z - z_0)$, analytic in the same domain as $f$. Hence

$$\lambda I - f(T) = (\lambda - f)(T) = (z_0 I - T)g(T) = g(T)(z_0 I - T).$$

Since $z_0 I - T$ is not invertible, so that $\lambda I - f(T)$ is not invertible, so that $\lambda \in \sigma(f(T))$. □

# Appendix A. Some topology facts

**Proposition 1** *A closed subset of a compact set is compact.*

**Proposition 2** *A subset of a complete set is complete if and only if it is closed.*

# Appendix B. Homework and exam sets

**Definition 1** *A topological space $X$ is* **locally compact** *if for all $x \in X$ there exists a compact set that contains a neighborhood of $x$.*

**Proposition 2** *A subset of $\mathbb{F}^n$ is closed and bounded if and only if it is compact.*

**Homework 1** **MATH 411 • Spring 2011**

**Due on Friday, 4 February:**

1. Let $X$ be a normed linear space. Prove that the functions $+ : X \times X \to X$ and $\cdot : \mathbb{F} \times X \to X$ are continuous. (Use the product or the product metric topology on the product of (metric) spaces.)
2. Let $X \subseteq \mathbb{R}$, and let $f_n : X \to \mathbb{F}$ be differentiable functions such that $f_n \to g$ uniformly and $f_n' \to h$ uniformly. Prove that $g$ is a differentiable function and that $g' = h$. (Hint: Apply the Mean Value Theorem to $f_n - f_m$.)
3. Let $X$ be a topological space, and let $S_1 \supseteq S_2 \supseteq \cdots$ be a nested sequence of closed compact subsets. Prove that $\cap_n S_n = \emptyset$ if and only if some $S_n$ is empty.
4. Let $a, b \in \mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be a bounded measurable function. Prove that

$$\lim_{p \to \infty} ||f||_p \leq \inf\{\sup\{|f(t)| \ : \ t \in E\} \ : \ E \subseteq [a, b], m^*([a, b] \setminus E) = 0\}.$$

The quantity on the right in the display is called the $\mathbb{L}^\infty$-norm.

**Correction**

Problem 3 above was originally phrased as follows: Let $X$ be a topological space, and let $S_1 \supseteq S_2 \supseteq \cdots$ be a nested sequence of compact subsets. Prove that $\cap_n S_n = \emptyset$ if and only if some $S_n$ is empty.

Here is a counterexample: Let $X = \mathbb{R}$. The open sets are $\mathbb{R}, [0, r),\ r \in \mathbb{R}$. This forms a topology. The sets $S_n = (0, \frac{1}{n}]$ are compact and non-empty, yet $\cap S_n = \emptyset$.

**Homework 2** **MATH 411 • Spring 2011**

**Due on Friday, 11 February:**

1. Let $\Sigma$ be a $\sigma$-algebra on a set $X$. A function $\mu : \Sigma \to \mathbb{R} \cup \{\pm\infty\}$ is called a **signed measure** if $\mu(\emptyset) = 0$ and if $\mu$ is countable additive, i.e., if for any pairwise disjoint $A_1, A_2, \ldots$ in $\Sigma$, $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$. Prove that the range of $\mu$ contains at most one value of $\{-\infty, \infty\}$.

2. Let $\Sigma$ be a $\sigma$-algebra on a set $X$. A function $\mu : \Sigma \to \mathbb{C}$ is called a **complex measure** if $\mu(\emptyset) = 0$ and if $\mu$ is countable additive, i.e., if for any pairwise disjoint $A_1, A_2, \ldots$ in $\Sigma$, $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$. Prove that $Im\,\mu$, $Re\,\mu$ are signed measures.

3. Let $\mu$ be a signed measure on a $\sigma$-algebra $\Sigma$ on a set $X$. This exercise will prove the **Hahn-Jordan decomposition theorem**: there exist $P$ and $N$ in $\Sigma$ such that $X = P \cup N$, $P \cap N = \emptyset$, and for all $E \in \Sigma$, $\mu(E \cap P) \geq 0$ and $\mu(E \cap N) \leq 0$. If $\mu_P, \mu_N : \Sigma \to [0, \infty]$ are defined as $\mu_P(E) = \mu(E \cap P)$ and $\mu_N(E) = -\mu(E \cap N)$, then $\mu_P, \mu_N$ are measures, and $\mu = \mu_P - \mu_N$.

   (i) Let $\mathbb{P}$ (resp. $\mathbb{N}$) be the set of all sets $A \in \Sigma$ such that for all $E \in \Sigma$, $\mu(E \cap A) \geq 0$ (resp. $\mu(E \cap A) \leq 0$). Prove that $\mathbb{P}$ and $\mathbb{N}$ are not empty, and that they are closed under countable unions.

   (ii) By Exercise 2.1, $\mu$ never takes on either the value $\infty$ or $-\infty$. We will assume in the sequel that $\mu$ does not have $\infty$ in its range. Let $B = \sup\{\mu(A) : A \in \mathbb{P}\}$. Prove that there exist $A_1 \subseteq A_2 \subseteq \cdots$ in $\mathbb{P}$ such that $B = \lim_{n\to\infty} \mu(A_n)$. Prove that $P = \cup_n A_n \in \mathbb{P}$ and that $B < \infty$.

   (iii) Set $N = X \setminus P$. Prove that $N \in \mathbb{N}$. (Suppose not. Then there exists $E \subseteq N$ such that $E \in \Sigma$ and $\mu(E) > 0$. Necessarily (why?) $E \notin \mathbb{P}$. Thus there exists $F \subseteq E$ such that $F \in \Sigma$ and $\mu(F) < 0$. Choose the smallest possible positive integer $n_1$ such that for some such $F$, call it $F_1$, $\mu(F_1) < -1/n_1$. Necessarily (why?) $E \setminus F_1 \notin \mathbb{P}$. In general, let $n_k$ be the smallest positive integer such that for some $F_k \subseteq E \setminus (F_1 \cup \cdots \cup F_{k-1})$, $F_k \in \Sigma$ and $\mu(F_k) < -1/n_k$. Set $F = \cup_k F_k$. Prove that $\mu(F) < -\sum_k 1/n_k \in (-\infty, 0)$, so that $\lim_k n_k = \infty$. Use the "greediness" of the $n_k$ to show that for all $G \in \Sigma$ with $G \subseteq E \setminus F$, $\mu(G) \geq 0$. Conclude that $E \setminus F \in \mathbb{P}$. Prove that $\mu(E \setminus F) = 0$ and that $\mu(E \setminus F) > 0$.)

   (iv) Prove that $\mu_P, \mu_N$ as defined in the introduction are measures.

   (v) Prove that $\mu = \mu_P - \mu_N$.

**Homework 3** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Friday, 18 February:**

1. Let $(X, \Sigma, \mu)$ be a measure space. Prove that $\mathbb{L}^\infty$ is complete.
2. Let $X$ be a non-empty set with a $\sigma$-algebra $\Sigma$ on it. Let $M = M(X, \Sigma)$ be the set of all $\mathbb{F}$-valued measures $\mu$ on $\Sigma$ for which $\sup\{|\mu(A)| : A \in \Sigma\} < \infty$.
   - (i) Prove that $M$ is a vector space over $\mathbb{F}$ under the obvious operations.
   - (ii) Define $\| \ \|$ on $M$ by $\|\mu\| = \sup\{|\mu(A)| : A \in \Sigma\}$. Prove that $\| \ \|$ is a norm on $M$.
   - (iii) Prove that $M$ is complete in the norm, so that $M$ is a Banach space. (Hint: Before you read the rest of my hint, be aware that my solution may not be the slickest one. Let $\{\mu_n\}$ be a Cauchy sequence. Find $\mu$ that is a pointwise limit. If $A_1, A_2, \ldots$ are pairwise disjoint in $\Sigma$, prove that $\{\sum_{i=1}^n \mu(A_i)\}_n$ is a Cauchy sequence in $\mathbb{F}$.)
3. Let $V$ be a non-trivial vector space over $\mathbb{F}$. The goal of this exercise is to prove that $V$ has a **basis**, i.e., $V$ has a subset $B$ such that every element of $V$ can be written uniquely as a finite linear combination of the elements of $B$. We can only prove this if we allow the axiom of choice, or one of its equivalent formulations. We will take the following equivalent formulation of the axiom of choice: **Zorn's lemma:** Let $S$ be a non-empty set with partial order on its elements. Suppose that for any totally ordered subset $C$ of $S$ there exists $B \in S$ such that for all $c \in C$, $c \leq B$. (This is phrased as: **every chain in $S$ has an upper bound**.) Then there exists $M \in S$ such that for all $s \in S$, either $s$ and $M$ are incomparable or $s \leq M$. (This is phrased as: $S$ **has a maximal element**.)

   With that, prove that $V$ has a basis as follows. Let $S$ be the collection of all linearly independent* subsets of $V$. Prove that $S$ is not empty, impose the natural partial order on $S$, prove that every chain in $S$ has an upper bound, invoke Zorn's lemma, and then prove that the maximal element in $S$ is a basis.

---

* Definition of **linear independence**: a (possibly infinite) set $L$ is linearly independent if any finite linear combination of its elements being zero implies that all the coefficients are zero.

**Homework 4** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Friday, 25 February:**

1. Let $X$ be a normed vector space with norm $\| \ \|$. The series $\sum_{n=1}^{\infty} x_n$ **converges** if the sequence $\{\sum_{i=1}^{n} x_i\}_n$ converges in $X$ (in the norm, of course). The series $\sum_{n=1}^{\infty} x_n$ **converges absolutely** if $\sum_{n=1}^{\infty} \|x_n\|$ converges. Prove that $X$ is complete if and only if every absolutely convergent series in $X$ converges.

2. Let $p \in (0,1)$, let $a, b \in \mathbb{R}$ with $a < b$, and let $m$ denote the Lebesgue measure. Let $X$ be the set of all $m$-equivalence classes of Lebesgue-measurable functions $f : [a,b] \to \mathbb{F}$. Define
$$((f))_p = \left( \int_a^b |f|^p dm \right)^{1/p}.$$
Let $Y$ be the subset of $X$ consisting of those $f$ for which $((f))_p < \infty$. Prove that $((\ ))_p$ is not a norm.

3. Let $X$ and $Y$ be vector spaces, and let $T : X \to Y$ be linear. Define
$$N(T) = \textbf{nullspace of } T = \textbf{kernel of } T = \{x \in X : T(x) = 0\}.$$
For the rest of this exercise, assume that $X$ and $Y$ are both normed.
   (i) Give an example of $X, Y$ and $T$ such that $N(T)$ is not closed.
   (ii) Now assume that $T \in B(X,Y)$. Prove that $N(T)$ is a closed subspace of $X$. Give an example of $X, Y, T$ such that the range of $T$ is not a closed subset of $Y$.

4. Let $X$ be a normed vector space and $M$ a closed subspace. In class we defined the obvious norm on $X/M$ under the condition that $X$ is also complete. Is this complete assumption necessary? (The book seems to say on page 71 that it is not necessary.)

## Addition

It is indeed not necessary. The corrected class notes now reflect that.

**Homework 5** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Friday, 4 March:**

1. Let $X = \mathbb{F}^n$. Prove that $X^*$ is isomorphic as a ((complete) normed) vector space to $X$.
2. Let $X$ be a normed vector space.
   (i) Let $S \subseteq X$. Prove that
   $$S^\circ = \{f \in X^* : f(s) = 0 \text{ for all } s \in S\}$$
   is a closed linear subspace of $X^*$.
   (ii) Let $T \subseteq X^*$. Prove that
   $$^\circ T = \{x \in X : f(x) = 0 \text{ for all } f \in T\}$$
   is a closed linear subspace of $X$.
   (iii) Let $M$ be a closed linear subspace of $X$. Prove that $^\circ(M^\circ) = M$.
3. Let $V$ be a normed vector space over $\mathbb{C}$. Prove that the function
   $$\varphi : B_{\mathbb{C}}(V, \mathbb{C}) \to B_{\mathbb{R}}(V, \mathbb{R})$$
   given by $\varphi(f) = Re(f)$ is an isometric isomorphism.
4. Let $V$ be a vector space over $\mathbb{C}$. Let $B$ be a basis of $V$. Prove that $B \cup iB$ is a basis of $V$ as a vector space over $\mathbb{R}$.

**Homework 6** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Friday, 18 March:**

1. Let $X$ be the set of all bounded functions $\mathbb{R} \to \mathbb{F}$.
   (i) Prove that $X$ is an $\mathbb{F}$-vector space.
   (ii) Define $\| \ \| : X \to \mathbb{R}$ by $\|f\| = \sup\{|f(x)| : x \in \mathbb{R}\}$. Prove that $\| \ \|$ is a norm.
   (iii) Prove that $X$ is complete under the norm.
2. With $X$ as in the previous problem, prove that there exists $T \in X^*$ such that
   (i) $\|T\| = 1$.
   (ii) If $f \in X$ is real-valued, then $\inf\{f(x) : x \in \mathbb{R}\} \leq T(f) \leq \sup\{f(x) : x \in \mathbb{R}\}$.
   (iii) If $f \in X$ is in non-negative real-valued, then $T(f) \in [0, \infty)$.
   (iv) If $f, g \in X$ such that for some $r \in \mathbb{R}$, $f(x + r) = g(x)$ for all $x \in \mathbb{R}$, then $T(f) = T(g)$.
3. Suppose that $x \in \ell^\infty$ and that $\{\frac{x_1 + \cdots + x_n}{n}\}$ converges. Is the Banach limit of such $x$ uniquely determined? Prove or find a counterexample. (I don't know an answer.)
4. Let $X, Y$ be normed vector spaces $T \in B(X, Y)$ and $M = \operatorname{kernel} T$. Prove that $T$ is an open map if and only if the induced map $\tilde{T} \in B(X/M, Y)$ is a topological homeomorphism. (For $\tilde{T}$, see Remark 12.5.)
5. Let $X$ and $Y$ be Banach spaces, and let $T \in B(X, Y)$. Prove that the following are equivalent:
   (i) The range of $T$ is closed.
   (ii) There exists $c$ such that for all $x \in X$, $d(x, \operatorname{kernel} T) \leq c \|T(x)\|$.
   (iii) There exists $c'$ such that for any $y$ in the range of $T$, there exists $x \in X$ such that $T(x) = y$ and $\|x\| \leq c' \|y\|$.

**Homework 7** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Friday, 1 April:**

1. Let $e_n \in \ell^1$ be the sequence with 1 in the $n$th spot and 0 elsewhere. Clearly $e_1, e_2, \ldots$ are linearly independent.
   (i) Prove that for each $n$ there exists a vector space basis $B$ of $\ell^1$ that contains $e_1, e_2, \ldots$ and for which $Span\{e_1, \ldots, e_n\} \cap \overline{Span(B \setminus \{e_1, e_2, \ldots\})} = 0$.
   (ii) Find a vector space basis $B$ of $\ell^1$ that contains $e_1, e_2, \ldots$ and for which $e_1 \in \overline{Span(B \setminus \{e_1, e_2, \ldots\})}$.
   * (iii) Prove that for every vector space basis $B$ of $\ell^1$ that contains $e_1, e_2, \ldots$, there exists $n$ such that $Span\{e_1, \ldots, e_n\} \cap \overline{Span(B \setminus \{e_1, e_2, \ldots\})} \neq 0$.
2. Let $Y = \{x \in \ell^\infty : \{nx_n\} \in \ell^\infty\}$. Then $Y$ is a vector space. Define the norm on $Y$ as $\|x\| = \|x\|_\infty + \|\{nx_n\}\|_\infty$. Prove that $\|\ \|$ is a norm and that $Y$ is complete in this norm.
3. Let $M = \{\{a_n\} \in \ell^\infty : a_{2n} = 0 \text{ for all } n \in \mathbb{N}\}$, $N = \{\{b_n\} \in \ell^\infty : nb_{2n} = b_{2n-1} \text{ for all } n \in \mathbb{N}\}$, and set $X = M + N$. By Example 10.1 (12), $X$ is a non-closed subset of $\ell^\infty$, so that $X$ is not complete. Let $Y$ be as in the previous exercise. Define $T : X \to Y$ by $T(x) = \{x_n/n + x_{2n}\}_n$. Then $T$ is well-defined and it is certainly linear. Prove that $T$ is surjective but not continuous.
4. Let $B$ be a vector space basis of $c_0$ that contains the canonical elements $e_1, e_2, \ldots$. Define $T : c_0 \to c_0$ by $T(e_n) = e_n/n$, $T(b) = b$ for all $b \in B \setminus \{e_1, e_2, \ldots\}$, and extend $T$ linearly to all of $X$. Prove that for some $B$, $T$ is not continuous. (Hint: Define $b_n = (0, 0, \ldots, 0, 1, 1/n, 1/n^2, 1/n^3, \ldots)$ (the first $n-1$ entries are 0). Prove that the set containing all the $e_n$ and all the $b_n$ is linearly independent. (I can do this with a Vandermonde matrix.) Prove that for $n > 1$, $-e_n + b_n$ has $\ell^\infty$-norm $1/n$ and $-e_n/n + b_n$ has $\ell^\infty$-norm $(n-1)/n$.)

**Homework 8** **MATH 411 • Spring 2011**

**Make sure you are in class each day on April 11, 13, 15, from 2:10pm until 3:00pm.**

**Solve two problems of your choice. Due on Friday, 8 April:**

1. Let $X$ be a Banach space. Prove that a subset $S$ of $X^*$ is bounded if and only if for each $x \in X$, $\{|f(x)| : f \in S\}$ is bounded. (Hint: this is easy.)
2. Does the conclusion of the previous exercise hold if $X$ is not complete in its norm? Give a proof or a counterexample.
3. Let $A$ be any subset of a Hilbert space $H$. Prove that the closure of the linear span of $A$ equals $(A^\perp)^\perp$.

**Homework 9** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Friday, 22 April:**

1. Let $V$ be a vector space. Prove that any two vector space bases of $V$ have the same cardinality. If $V$ is a Hilbert space, prove that any two (Hilbert space) bases of $V$ have the same cardinality.
2. Let $H$ be a Hilbert space. Prove that $H$ has a countable (Hilbert space) basis if and only if $H$ has a countable dense subset.
3. Prove that two Hilbert spaces are isomorphic if and only if they have (Hilbert space) bases of the same cardinality. (Recall: An isomorphism of Hilbert spaces is a vector space isomorphism that preserves inner products.)
4. Let $X$ and $Y$ be Hilbert spaces and $T \in B(X, Y)$. Prove that $\|T\| = \|T^*\| = \|T^* \circ T\|^{1/2}$.
5. Let $X$ be a Banach space in which the Parallelogram Law holds. Prove that $X$ is a Hilbert space.

**Homework 10** **MATH 411 • Spring 2011**

**Solve two problems of your choice. Due on Wednesday of the thesis week at noon.**

1. Determine the spectrum of the unilateral shift $T : \ell^2 \to \ell^2$. (See Example 20.4, Corollary 20.7, Corollary 20.9.)
2. Let $X$ be a Banach space and let $T \in B(X)$. Prove that $\lim_{n\to\infty} \|T^n\|^{1/n} = \inf_n \|T^n\|^{1/n}$. (Hint/trick: For any $k \in \mathbb{N}$, write $n = q_n k + r_n$ with $0 \le r_n < k$.)
3. Give an example of a Banach space $X$ and an invertible $T \in B(X)$ such that $\|T\| \left\|T^{-1}\right\| \neq 1$.
4. Let $X$ be a Banach space, and let $S, T \in B(X)$. (They need not commute.) Prove that $\sigma(S \circ T) \setminus \{0\} = \sigma(T \circ S) \setminus \{0\}$. (Hint: if $S \circ T - \lambda I$ has an inverse $U$, fashion an inverse of $T \circ S - \lambda I$ out of $U, S, T, \lambda, I$ in a clever way.)

**Exam 1** **MATH 411 • Spring 2011**

You may take three hours for this exam, with at most one 10-minute break in between. You are not allowed to use any notes, books, people. If you have questions for me, you may call me in my office 503 517 7399 or at home 503 788 6084 (after 6:45am and before 10pm). If your time is up, but you still have ideas, you may keep going, but mark how much was done after the allowed time and how much extra time you took. Do not go back to the exam after a break after your allotted time was up.

**Below you will find five theorems and five proofs. Match the theorems with their proofs. If a theorem does not have a proof, write the missing proof; and if a proof does not have a statement, write the missing theorem. Correct and complete any given theorems and proofs as necessary.**

**Do not allow a domino effect of errors.**

A. Theorem: Let $X$ be a normed vector space, and let $e_1, \ldots, e_n$ be elements of norm 1. Let $T \in X^*$. Then $\|T\| \geq \max\{|T(e_1)|, \ldots, |T(e_n)|\}$. If $X$ is spanned by $e_1, \ldots, e_n$, and if $X$ has norm determined by __________________________________,
__________________________________________________________,
then $\|T\| = \max\{|T(e_1)|, \ldots, |T(e_n)|\}$.

B. Theorem: Let $X, Y$ be a normed vector spaces, and let $T \in B(X, Y)$. Define $U : Y^* \to X^*$ by $U(f) = f \circ T$. Then $U \in B(Y^*, X^*)$ and $\|U\| \leq \|T\|$.

C. Theorem: Let $X, Y, Z$ be normed vector spaces, $T \in B(X, Y)$ and $U \in B(Y, Z)$. Then $U \circ T \in B(X, Z)$ and $\|U \circ T\| \leq \|U\| \, \|T\|$.

D. Theorem: Let $X, Y$ be normed vector spaces, and $T \in B(X, Y)$. Let $U : Y^* \to X^*$ be defined as in Theorem B. Then $(\text{range } T)^\circ = \text{kernel}\, U$ and $\text{kernel}\, T = {}^\circ(\text{range } U)$.

E. Theorem: Let $Z$ be a subset of a normed vector space $X$, and let $Y$ be the closed linear span of $Z$. Then $Z^\circ = Y^\circ$, and ${}^\circ(Z^\circ) = Y$.

---

Recall that for $S \subseteq X$, $S^\circ = \{f \in X^* : f(s) = 0 \text{ for all } s \in S\}$, and that for $T \subseteq X^*$, ${}^\circ T = \{x \in X : f(x) = 0 \text{ for all } f \in T\}$.

(i) Proof: For all non-zero $x \in X$, $\|x\| = \left\|T^{-1}T(x)\right\| \leq \left\|T^{-1}\right\| \|T\| \|x\|$, so that $1 \leq \left\|T^{-1}\right\| \|T\|$, whence $\left\|T^{-1}\right\| \geq 1/\|T\|$. An inequality is possible: [Fill in the rest.]

(ii) Proof: If $f \in Y^*$, then $f \in \operatorname{kernel} U$ if and only if $U(f) = 0$, which holds if and only if $f \circ T = 0$, which holds if and only if for all $x \in X$, $f \circ T(x) = 0$, which holds if and only if $f \in (\text{range } T)^\circ$. This proves that ______________________________. [Fill in the rest.]

(iii) Proof: Let $f, g \in Y^*$, $r \in \mathbb{F}$. Then for all $x \in X$, $U(f + rg)(x) = (f + rg) \circ T(x) = f \circ T(x) + rg \circ T(x) = U(f)(x) + rU(g)(x) = (U(f) + rU(g))(x)$, so that ______________________________________________________. Furthermore, by ____________________ we have that ____________________________________________ _______________________________________________________________________.

(iv) Proof: If $f \in X^*$, then $f \in Z^\circ$ if and only if $Z \subseteq \operatorname{kernel} f$. Since the kernel of $f$ is a closed subspace, $Z \subseteq \operatorname{kernel} f$ if and only if $Y \subseteq \operatorname{kernel} f$. Thus _________________________________________. This implies that ${}^\circ(Z^\circ) = {}^\circ(Y^\circ)$, and by Homework 5.2.(iii), since $Y$ is closed, ${}^\circ(Z^\circ) = Y$.

(v) Proof: Since $\operatorname{kernel} U = (\text{range } T)^\circ$, it follows that ${}^\circ(\operatorname{kernel} U) = {}^\circ((\text{range } T)^\circ) = \overline{\text{range } T}$ by _____________________________________________________________.

**Exam 2** **MATH 411 • Spring 2011**

The 17 students formed 6 groups, and each group came up with 2 problems. The following is the list of all exam questions.

1. Let $H$ be a Hilbert space and let $A$ be a continuous linear operator from $H$ to itself. Using the Riesz Representation Theorem one is able to prove that there exists a unique linear operator $A^*$ such that for any $a, b \in H$, $\langle Aa, b\rangle = \langle a, A^*b\rangle$. This operator is known as the adjoint of $A$.

   Given the above, prove three of the following five statements.

   (i) $A^{**} = A$.

   (ii) If $A$ invertible then $A^*$ is invertible, and $(A^*)^{-1} = (A^{-1})^*$.

   (iii) $(A+B)^* = A^* + B^*$.

   (iv) $(\lambda A)^* = \overline{\lambda} A^*$, where $\overline{\lambda}$ denotes the complex conjugate of $\lambda$.

   (v) $(AB)^* = B^*A^*$.

   Then prove that $\text{Ker}(A^*) = \text{Im}(A)^{\perp}$. From this it follows immediately (with a result you proved in homework) that $\text{Ker}(A^*)^{\perp} = \overline{\text{Im}(A)}$.

2. (i) State the Uniform Boundedness Principle.

   (ii) Let $X$ be a Banach space and $Y$ be a normed vector space. Prove that a set $S \subseteq B(X, Y)$ is bounded if and only if for all $x \in X$ and all $g \in Y^*$, $\{g \circ T(x) : T \in S\}$ is bounded in $\mathbb{F}$.

   (iii) Show that in the above proposition, the assumption that $X$ is Banach is necessary. (*Hint*: In constructing a counter-example, you may start with $X = \oplus_{n=1}^{\infty} \mathbb{F}$.)

3. (i) State the Uniform Boundedness Principle.

   (ii) Are all the assumptions necessary? If not, prove it without each unnecessary assumption. If yes, give counterexamples to show that the principle fails when one relaxes the following assumptions:
   T he domain is Banach.
   T he functions are continuous.
   T he functions are linear.

4. What is wrong with the following version of the Open Mapping Theorem. Explain why this version fails:

   **Theorem.** *Let $X$ and $Y$ be normed vector spaces and let $T \in B(X, Y)$ be surjective. Then $T$ is an open mapping.*

5. Let $H$ be a Hilbert space and $T \in H^*$. Let $M = \text{Ker}\, T$. Prove that either $M^{\perp}$ is a vector space of dimension 1 of $M = H$.

6. Let X and Y be topological spaces. A function $f : X \to Y$ is continuous if for every

open set V in Y, $f^{-1}(V) = \{x \in X : f(x) \in V\}$ is open in X.

**If f : X → Y is continuous, then for any x ∈ X and any sequence in X converging to x, the composition of f with this sequence converges to f(x).**

(i) Prove: The converse of the bolded statement holds if one uses nets (instead of sequences). (Hints: Try the contrapositive. Definition: A function $f : X \to Y$ is continuous at a point $x \in X$ if given an open neighborhood V of f(x), there exists an open set $U \subseteq X$ such that $U \subseteq f^{-1}(V)$. Define the set $A = \{W : W \text{ is an open neighborhood of } x\}$, and then order it by reverse inclusion (so $w_1 \leq w_2$ if $w_1 \supseteq w_2$).)

(ii) Find a directed set $(A, \leq)$ with the following properties:

(a) A is not finite

(b) There exists exactly one element $a_0$ of A such that, for every (Hausdorff) topological space X, all convergent nets $S : A \to X$ must necessarily converge to $S_{a_0} \in X$.

7. Let $X$ be a normed linear space, and let $\varphi : X \to X^{**}$ be given by

$$(\varphi(x))(f) = f(x).$$

We say that $X$ is **reflexive** if $X^{**} = \{\varphi(x) : x \in X\}$, i.e., if $\varphi$ is an isomorphism. Prove that $X^*$ is reflexive if $X$ is reflexive.

8. Let $X$ be a normed linear space, and let $\varphi : X \to X^{**}$ be given by

$$(\varphi(x))(f) = f(x).$$

We say that $X$ is **reflexive** if $X^{**} = \{\varphi(x) : x \in X\}$, i.e., if $\varphi$ is an isomorphism. Suppose that $X$ is complete and that $X^*$ is reflexive. Prove that $X$ is reflexive.

9. Given any partially ordered set $(A, \leq)$, a subset $B$ of $A$ is a **cofinal subset** if for every $a \in A$ there exists $b \in B$ such that $a \leq b$. Given a function $f : X \to A$, we say that $f$ is a **cofinal function** if $f(X)$ is a cofinal subset of $A$. Let $A$ and $B$ be directed sets. Given two nets $(x_\alpha)$ and $(y_\beta)$ from $A$ to $B$, respectively, define $(y_\beta)$ to be a **subnet** of $(x_\alpha)$ if there exists a monotone cofinal function $h : B \to A$ such that $y_\beta = x_{h(\beta)}$.

(i) Let $X$ be a compact topological space. Assume that the collection $\{C_i : i \in I\}$ of closed subsets of $X$ has the **finite intersection property**, i.e., $\cap_{j \in J} C_j \neq \emptyset$ for all finite subsets $J$ of $I$. Prove that $\cap_{i \in I} C_i \neq \emptyset$.

(ii) Let $A$ be a directed set and $(x_\alpha)$ a net in (compact) $X$. For every $\alpha \in A$ define $E_\alpha = \{x_\beta : \beta \geq \alpha\}$. Prove that $E = \cap_{\alpha \in A} \overline{E_\alpha} \neq \emptyset$.

(iii) Let $x \in E$. Then each neighborhood $U$ of $x$ has a non-empty intersection with all $E_\alpha$. Use the fact above to construct a subnet $(y_\beta)$ of $(x_\alpha)$ that converges to $x$.

10. Define the **Chebyshev polynomials of the first kind** via the recurrence:

$$
\begin{aligned}
T_0(x) &= 1, \\
T_1(x) &= x, \\
T_x(n) &= 2xT_{n-1}(x) - T_{n-2}(x).
\end{aligned}
$$

  (i) Prove that for all integers $n \geq 0$, $T_n(\cos(\theta)) = \cos(n\theta)$.

  (ii) Prove that $\{T_n \circ \cos : n \geq 0\}$ is an orthogonal set in $\mathbb{L}^2([-1,1])$ (with the Lebesgue measure).

11. Give an example of a linear mapping between two normed vector spaces that is almost open but not open.

12. Let $\{a_n\}$ be a sequence in $\mathbb{F}$. Prove that $\sum_{n=1}^{\infty} a_n$ converges absolutely if and only if $\sum_{n\in\mathbb{N}} a_n$ converges in the net.

**Final exam** **MATH 411 • Spring 2011**

**Solve three of the five problems, in three contiguous hours. When solving a problem, you may use the results of preceding problems in this exam. You may refer to your class notes and to my notes on the web, but you may use no other sources. Due on Thursday at noon.**

Recall that an operator $T : X \to X$ is **compact** if the image of $T(B(0,1))$ is contained in a compact subset of $X$.

1. Let $X$ be a Hilbert space, and let $T \in B(X)$. Prove that $T$ is compact if and only if there exists a sequence $\{T_n\}$ of compact operators such that their ranges are finite-dimensional and such that $\|T - T_n\| \to 0$.
2. Let $X$ be a Hilbert space, and let $T$ be a compact operator on $X$. Prove that the closure of the range of $T$ is a separable Hilbert space, i.e., that it has a countable Hilbert space basis.
3. Let $a < b$ be real numbers. Consider $\mathbb{L}^2([a,b])$ (of complex-valued functions) as a Hilbert space. Let $k : [a,b] \times [a,b] \to \mathbb{C}$ be measurable, such that $|k|^2$ has a finite integral. Define $K : \mathbb{L}^2([a,b]) \to \mathbb{L}^2([a,b])$ by

$$(Kf)(x) = \int_a^b k(x,y) f(y) dy.$$

   (i) Prove that $K$ is a continuous linear transformation of vector spaces.

   (ii) Let $S$ be a Hilbert space basis of $\mathbb{L}^2([a,b])$. Let $g, h \in S$, and suppose that $k(x,y) = g(x)h(y)$. Prove that $K$ is compact. (You may assume Fubini's Theorem.)

   (iii) Assume that $k$ can be approximated arbitrarily closely in the $\mathbb{L}^2([a,b] \times [a,b])$-norm by a function of the form $\sum_{i=1}^r g_i(x)h_i(y)$, with $g_i, h_i \in \mathbb{L}^2([a,b])$. (This is indeed always true.) Prove that $K$ is compact.

4. Find the norm and at least two eigenvalues of the operator $K \in B_0(\mathbb{L}^2[0,3])$, given by

$$(Kf)(x) = \int_0^3 \chi_{[0,2]}(x) \chi_{[1,3]}(y) f(y) dy,$$

   where $\chi_S$ denotes the characteristic function of $S$. Prove everything.

5. Let $X$ be an infinite-dimensional Hilbert space. Let $T \in B(X)$ be compact. Prove that $0 \in \sigma(T)$.

**Solutions for the final exam**

Solution of 1: ($\Longrightarrow$) Let $T$ be compact. If $T = 0$, we may take all $T_n$ to be 0. So we assume that $T \neq 0$. For each $n$, let $U_n = \{B(y, 1/n) : \|y\| \leq \|T\|\}$. Then for all $x \in B(0,1)$, $T(x)$ has norm at most $\|T\|$, and so $U_n$ is a cover of $\overline{T(B(0,1/n))}$. Since $T$ is compact, there exists a finite subcover $U' = \{B(y_{n1}, 1/n), \ldots, B(y_{nk_n}, 1/n)\}$. Let $M_n = \mathrm{Span}\{y_{n1}, \ldots, y_{nk_n}\}$, and define $T_n = P_{M_n} \circ T$. Then the range of $T_n$ is finite-dimensional, and therefore $T_n$ is compact. Let $x \in B(0,1)$. Choose large $n$ and $y_{ni}$ such that $T(x) \in B(y_{ni}, 1/n)$. Then

$$\begin{aligned}
\|(T - T_n)(x)\| &\leq \|T(x) - y_{ni}\| + \|y_{ni} - T_n(x)\| \\
&< 1/n + \|P_{M_n}(y_{ni} - T(x))\| \\
&\leq 1/n + \|P_{M_n}\| \, \|(y_{ni} - T(x))\| \\
&< 2/n,
\end{aligned}$$

which proves that $T_n \to T$ in the norm.

($\Longleftarrow$) By possibly taking a subsequence of the $T_n$, we may assume that for all $n$, $\|T - T_n\| < 1/n$. Let $\{x_m\}_m$ be a sequence in $B(0,1)$. Since each $T_n$ is compact, there exists a subsequence $\{x_m^1\}_m$ of $\{x_m\}_m$ such that $\{T_1(x_m^1)\}_m$ converges, and then there exists a subsequence $\{x_m^2\}_m$ of $\{x_m^1)\}_m$ such that $\{T_2(x_m^2)\}_m$ converges, and then there exists a subsequence $\{x_m^3\}_m$ of $\{x_m^2)\}_m$ such that $\{T_3(x_m^3)\}_m$ converges, there exists a subsequence $\{x_m^1\}_m$ of $\{x_m\}_m$ such that $\{T_1(x_m^1)\}_m$ converges, etc. Let $y_n = x_n^n$. These $y_n$ yield a subsequence of $\{x_m\}$. Let $\epsilon > 0$. Choose $k$ such that $1/k < \epsilon/3$. Then there exists $N$ such that for all $m, n > N$, $\|T_k(y_m) - T_k(y_n)\| < \epsilon/3$. Then

$$\begin{aligned}
\|T(y_m) - T(y_n)\| &\leq \|T(y_m) - T_k(y_m)\| + \|T_k(y_m) - T_k(y_n)\| + \|T_k(y_n) - T(y_n)\| \\
&\leq 2\,\|T - T_k\| + \|T_k(y_m) - T_k(y_n)\| < \epsilon.
\end{aligned}$$

Thus $\{T(x_n)\}_n$ has a Cauchy and thus a convergent subsequence. It follows that every sequence in $\overline{T(B(0,1))}$ has a convergent subsequence. **We could stop here**, but perhaps you want to see a proof that every open cover of $\overline{T(B(0,1))}$ has a finite subcover. So, let $U$ be an open cover. By possibly adding more and smaller subsets to $U$, we may assume that each element of $U$ is of the form $B(x,r)$. Let $y_1 = 0 = T(0)$. Then there exists $B(x_1, r_1) \in U$ such that $y_1 \in B(x_1, r_1)$. Among all such $x_1, r_1$ choose one for which we can find the smallest positive integer $N_1$ with $1/N_1 \leq r_1$. If $\overline{T(B(0,1))} \subseteq B(x_1, r_1)$, we have found a finite subcover. Otherwise let $y_2 \in \overline{T(B(0,1))} \setminus B(x_1, r_1)$. There exists $B(x_2, r_2) \in U$ that contains $y_2$. Again, choose $x_2, r_2$ such that with smallest possible positive integer $N_2$ with $1/N_2 \leq r_2$. If $\overline{T(B(0,1))} \subseteq B(x_1, r_1) \cup B(x_2, r_2)$, we have found a finite subcover. Otherwise let $y_3 \in \overline{T(B(0,1))} \setminus \cup_{i=1}^2 B(x_i, r_i)$, etc. In this way we get a sequence $\{y_n\}$ in $\overline{T(B(0,1))}$ By what we have already proved, there exists a convergent subsequence $\{y_{n_k}\}_k$. Let $y$ be the limit. Then $y \in \overline{T(B(0,1))}$, so there exists $B(x,r) \in U$ that contains $y$. Let $M$ be such that for all $k > M$, $y_{n_k} \in B(x,r)$. Let $N$ be the smallest positive integer such that $1/N \leq r$. By the choice of the sequence, $N_{n_k} < N$. But then $r < 1/N_{n_k}$ for all $k$. It follows that $r < r_{n_k}$ for all $k$, whence the sequence $\{y_{n_k}\}_k$ is not Cauchy. This is a contradiction. So the construction of the $y_n$ must stop in finitely many steps, so that every open cover has a finite subcover.

Solution of 2: By Problem 1, there exists a sequence $\{T_n\}$ of compact operators with finite-dimensional ranges such that $T_n \to T$ in the norm. Without loss of generality for all $n$, $\|T_n - T\| < 1/n$. For each $n$ let $U_n$ be a basis of the range of $T_n$. Let $U = \cup_n U_n$. Then $U$ is a countable set. Let $x \in X$. Then for all $n$, $T(x) = T(x) - T_n(x) + T_n(x)$. Since $\|T(x) - T_n(x)\| < \|x\|/n$, it follows that $T(x)$ is within $\|x\|/n$ of the span of $U$. Thus $x \in \overline{\operatorname{Span} U}$. It follows that the range of $T$ and hence the closure of the range of $T$ are both in $\overline{\operatorname{Span} U}$. Clearly $\overline{\operatorname{Span} U}$ has a countable dense subset, and so the closure of the range of $T$ has a countable dense subset, and since this closure is a closed subset of a Hilbert space, it is a Hilbert space with a countable dense subset, i.e., it is separable.

Solution of 3: (i) is easy. For (ii), note that the range of $T$ is in the span of $g$, so that the range if finite-dimensional. Thus clealy $T$ is compact. For (iii), use Problem #1.

Solution of 4: For any $f \in \mathbb{L}^2([0,3])$, $K(f) = \int_0^3 \chi_{[0,2]}(x)\chi_{[1,3]}(y)f(y)dy = \chi_{[0,2]}(x)\int_1^3 f(y)dy$, so that

$$\begin{aligned}
\|K(f)\| &= \left(\int_0^3 \left|\chi_{[0,2]}(x)\int_1^3 f(y)dy\right|^2 dx\right)^{1/2} \\
&= \left(\int_0^2 \left|\int_1^3 f(y)dy\right|^2 dx\right)^{1/2} \\
&= \left(2\left|\int_1^3 f(y)dy\right|^2\right)^{1/2} \\
&= \sqrt{2}\left|\int_1^3 f(y)dy\right|.
\end{aligned}$$

By Hölder's inequality, as $f(y) = 1 \cdot f(y)$, this is at most $\sqrt{2}\sqrt{2}\,\|f\|_2$. Thus $\|K\| \leq 2$. Furthermore, if $f(y) = \chi_{[1,3]}$, then $\|f\| = \sqrt{2}$, $K(f)(x) = 2\chi_{[0,2]}(x)$, and $\|K(f)\| = 2\sqrt{2}$, so that $\|K\| = 2$. Also, by inspection, $\chi_{[0,1]}$ is an eigenvector with eigenvalue 0, and $\chi_{[0,2]}$ is an eigenvector with eigenvalue 1.

Solution of 5: Let $T_n$ be as in Problem 1. Suppose that $T$ is invertible. By a theorem from class (Theorem 20.6), the set of invertible operators on $X$ is open, so that since

$\|T - T_n\| \to 0$, for all large $n$, $T_n$ is invertible as well. But $T_n$ is not injective as its range is finite-dimensional. Thus $T$ cannot be invertible.

Solution of 5 (most frequent solution): Suppose that $T$ is invertible. Then by the Inverse Function Theorem, $T^{-1}$ is continuous. Thus $I = T \circ T^{-1}$ takes $B(0,1)$ to a compact subset of $X$. In particular, any infinite orthonormal subset of $X$ is taken by $I$ to a subset of a compact set, thus since orthonormal sets are discrete, any infinite orthonormal subset is taken by $I$ to a compact set. But then the unit balls around the elements of this orthornormal set is an open cover without a finite subcover, which gives a contradiction. Thus $T$ is not invertible, so that $0 \in \sigma(T)$.