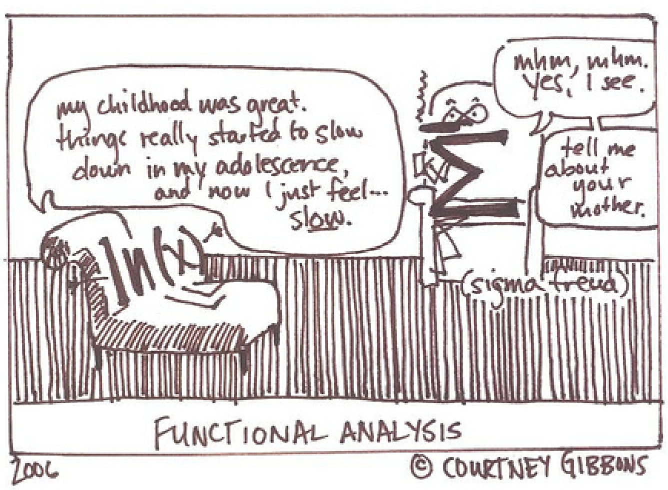
**M383C NOTES**

ARUN DEBRAY
SEPTEMBER 23, 2015

These notes were taken in UT Austin's Math 383C class in Fall 2015, taught by Todd Arbogast. I live-TeXed them using `vim`, and as such there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu.

CONTENTS






Source: http://brownsharpie.courtneygibbons.org/?p=78

Lecture 1.

## General Remarks: 8/26/15

Though the course name is "Methods of Applied Mathematics," this is a misnomer; the course is really about functional analysis.

The course will use the Canvas website (http://canvas.utexas.edu/), and office hours will be after class (modulo lunch), Mondays and Wednesdays from 12:30 to 1:50. Under UT Direct, there's also a CLIPS page, but that's less central to the course.

The textbook is a set of course notes; it hasn't changed much since 2013, so if you have that version, you'll be fine. They'll be ready at the copy center by Friday or Monday.

Homework will be due every week, assigned one Friday, and due the next. The first assignment will be due in a little over a week. We're encouraged to work in groups, but must write up our own individual proofs. Midterms will be weeks 7 and 12, probably, and will be topical; the final, at the end of the semester, will be comprehensive.

In this course, we'll cover chapters 2 – 5 of the lecture notes. Some elementary topology and Lesbegue integration (the first chapter) will be assumed.

Now, for some math. The professor is an applied mathematician, doing numerical analysis, and more specifically, approximation of differential equations. Functional analysis is useful for that, but also plenty of other fields, even including abstract algebra! Nonetheless, the course will be presented from an applied perspective.

The background is that we're trying to solve a problem of the form $T(u) = f$. Here, $T$ is a model or differential equation; it's some kind of operator. $f$ is the data that we're given, and we want to find the solution $u$. We use the framework of functional analysis to understand the nature of the functions $u$ and $f$: their properties and what classes of functions they live in. We also want to know the nature of the operator $T$. In particular, we'll focus on cases where $T$ is linear, since anything nonlinear can usually be locally approximated with a linear one. Thus, we should start with the linear case.

The set of all functions is a vector space, of course, so we're led to study vector spaces. At the undergraduate level, one studies finite-dimensional spaces, but here we'll use infinite-dimensional ones. Vector spaces also give us the required linearity. But since we also have questions of convergence, we'll introduce topology, so this course combines algebra and topology.

In this class, $\mathbb{F}$ will denote a field, either $\mathbb{R}$ or $\mathbb{C}$ (a lot of the time, the stuff we're doing won't depend on which).

**Definition.** Let $X$ be a vector space over $\mathbb{F}$. Then, $X$ is a *normed linear space* (henceforth NLS) if it has a *norm*, a function $\|\cdot\| : X \to \mathbb{R}^+ = [0,\infty)$ such that for every $x, y \in X$ and $\lambda \in \mathbb{F}$,

- $\|\lambda x\| = |\lambda|\|x\|$,
- $\|x\| = 0$ iff $x = 0$, and
- $\|x + y\| \leq \|x\| + \|y\|$.

The last stipulation is called the *triangle inequality*.

These conditions on the norm mean it's a measure of size: stretching a vector stretches the norm, the only thing with size 0 is the origin, and the triangle inequality corresponds to the familiar geometric one. It turns out these are the only properties we need to measure size.

**Example 1.1.**

(1) $d$-dimensional *Euclidean space* $\mathbb{F}^d$ comes with a familiar norm: if $x = (x_1, \dots, x_n)$ for $x_j \in \mathbb{F}$, then

$$\|x\| = \sqrt{\sum_{j=1}^{d} |x_j|^2}.$$

Sometimes, this is simply denoted $|x|$. Thus, whenever we talk about $\mathbb{F}^d$, we really mean $(\mathbb{F}^d, \|\cdot\|)$, the normed linear space.

(2) If $a < b$, where $a, b \in [-\infty, \infty]$, let $C([a,b])$ denote the space of continuous functions $f : [a,b] \to \mathbb{F}$ such that $\sup_{x\in[a,b]} |f(x)|$ is finite.[1] This is indeed a vector space; then, it turns to a normed linear space with the norm

$$\|f\| = \sup_{x\in[a,b]} |f(x)|.$$

Notice that the norm must be finite, which is satisfied here. The first two properties are clearly satisfied, and because the absolute value is a norm on $\mathbb{R}$, then the triangle equality is also satisfied.

[1] Recall that the *supremum* of a set is its least upper bound: for example, $\sup(0,1) = 1$, even though 1 isn't part of the set. This distinguishes the supremum from the maximum.

(3) We can pair $C([a,b])$ with a different norm $\|\cdot\|_{L^1}$, defined by

$$\|f\|_{L^1} = \int_a^b |f(x)|\,\mathrm{d}x.$$

The integral certainly exists, since $f$ is continuous, but it might be infinite; thus, we assume that $a$ and $b$ are finite, so $[a,b]$ is compact, and

$$\int_a^b |f(x)|\,\mathrm{d}x \le (b-a) \sup_{x\in[a,b]} |f(x)|,$$

so we're bounded. It's also not that hard to show that $\|\cdot\|_{L^1}$ is a norm, as the integral is linear.

We now have two norms on $C([a,b])$; are they "the same?" Though the underlying vector spaces are the same, the measures of size are different, so as normed linear spaces they are not the same.

We can find more examples sitting inside other NLSes.

**Proposition 1.2.** *Let $(X,\|\cdot\|)$ be an NLS and $V\subseteq X$ be a linear subspace. Then, $(V,\|\cdot\|)$ is an NLS.*

It's easy to check that the three requirements are still met.

We can measure size, so since we're in a vector space, we can measure distance. In general, we have a metric. Specifially, if $(X,\|\cdot\|)$ is an NLS, define $d : X\times X\to\mathbb{R}^+$ by $d(x,y)=\|x-y\|$. Why is this a metric? It has to satisfy the following three properties for all $x,y,z\in X$.

(1) $d(x,y)=0$ iff $x=y$.
(2) $d(x,y)=d(y,x)$.
(3) $d(x,y)+d(y,z)\ge d(x,z)$.

It's easy to check that the $d$ induced from the norm is indeed a metric; each metric property follows from one of the norm properties.

And now that we can measure distance, we have a topology; specifically a metric topology, the simplest of all topologies. That is, a normed linear space is a metric space. To be specific, define the *ball of radius $r$ about $x$*, where $r>0$ and $x\in X$, is

$$B_r(x)=\{y\in X \mid d(x,y)<r\}.$$

This is an open ball, so the distance must be strictly less than $r$.

The topology is defined by setting $U\subseteq X$ to be *open* if for every $x\in U$, there exists an $r>0$ such that $B_r(x)\subseteq U$. In other words, an open set doesn't contain its boundary. A set $F\subseteq X$ is *closed* if the complement $F^c = X\setminus F$ is open.

**Definition.** A subset $F$ of a metric space $X$ is *sequentially closed* if whenever $\{x_n\}_{n=1}^\infty$ is a sequence in $F$ converging to an $x\in X$ (in the sense of the metric, i.e. $d(x_n,x)\to 0$), then $x\in F$.

In a metric space (this is *not* true in general!), $F$ is closed iff $F$ is sequentially closed.

Now, we have algebra (the vector space), the metric (giving us convergence, compactness, etc.), and the norm. How are they related?

**Proposition 1.3.** *In an NLS $X$, addition, scalar multiplication, and the norm are all continuous functions.*

*Proof.* We'll prove this for addition and the norm; scalar multiplication is analogous to addition.

Addition is a function $+ : X\times X\to X$. Let $\{x_n\}\subseteq X$ with $x_n\to x$ and $\{y_n\}\subseteq X$ with $y_n\to y$. Continuity is equivalent to $\{x_n+y_n\}\to x+y$ for all such sequences. That is, I need $d(x_n+y_n, x+y)\to 0$, but that's equivalent to $\|(x_n+y_n)-(x+y)\|\to 0$.

Since $x_n\to x$ and $y_n\to y$, then $\|x_n-x\|\to 0$ and $\|y_n-y\|\to 0$. It looks like we should use the triangle inequality.

$$\begin{aligned}\|(x_n+y_n)-(x+y)\| &= \|(x_n-x)+(y_n-y)\|\\ &\le \|x_n-x\|+\|y_n-y\|\to 0.\end{aligned}$$

The norm is a little different. Suppose $x_n\to x$, which means we need to show that $\|x_n\|\to\|x\|$. Well,

$$\begin{aligned}\|x\| &= \|x-x_n+x_n\|\\ &\le \|x-x_n\|+\|x_n\|\\ &\le 2\|x-x_n\|+\|x\|.\end{aligned}$$

Since we've sandwiched $\|x - x_n\|$, then $\lim\|x_n\| = \|x\|$.[2] ⊠

Lecture 2.

## Banach Spaces: 8/28/15

Recall that if $(X, \|\cdot\|)$ is an NLS, we have a metric $d(x, y) = \|x - y\|$ and a topology. More generally, if $(X, d)$ is a metric space, $x_n \to x$ is the same as $d(x_n, x) \to 0$. In our case, this means that $\|x_n - x\| \to 0$.

**Definition.** A sequence $\{x_n\}_{n=1}^{\infty}$ is a *Cauchy sequence* if $\lim\limits_{n,m\to\infty} d(x_n, x_m) = 0$.

Here, $n$ and $m$ go to infinity independently, which might be confusing; an alternate way to phrase this is that $\{x_n\}$ is Cauchy if for all $\varepsilon > 0$, there exists an $N = N_\varepsilon > 0$ such that $d(x_n, x_m) \le \varepsilon$ whenever $m, n \ge N$.

In a Cauchy sequence, the terms get closer and closer together, but do they converge? Consider $(0, \infty)$ and $x_n = 1/n$. This is Cauchy, but would converge to 0, which isn't part of our set; in a sense, it's a "hole" in our set. This is annoying.

**Definition.**

- A metric space $X$ is *complete* if every Cauchy sequence on $X$ converges in $X$.
- A complete NLS is called a *Banach space*.

We'll also give some properties of subspaces of NLSes.

**Definition.** Let $X$ be an NLS. A set $M \subseteq X$ is *bounded* if there exists an $R > 0$ such that $M \subseteq \overline{B_R(0)} = \{x : \|x\| \le R\}$.

Equivalently, $M$ is bounded if there's a finite $R$ such that $\|x\| \le R$ for all $x \in M$.

**Proposition 2.1.** *Every Cauchy sequence in an NLS is bounded.*

*Proof.* The idea is that all but a finite number of points in a sequence are within distance 1 of each other.

Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in an NLS $X$. By definition (using $\varepsilon = 1$), there's an $N > 0$ such that $\|x_n - X_N\| \le 1$ for all $n \ge N$. Using the triangle inequality, $\|x_n\| \le \|x_N\| + 1$ for all $n \ge N$.

Now, let $M = \max\{\|x_1\|, \dots, \|x_{N-1}\|\}$ and $R = \max\{\|x_N\| + 1, M\}$; both of these are finite sets, and therefore have maxima. Thus, $\|x_n\| \le R$ for all $n$. ⊠

Even if the limit isn't there, the sequence is still bounded, which is nice. Also, notice how we used the norm; boundedness in metric spaces maybe isn't so interesting.

**Example 2.2.** Let's give some examples of Banach spaces.

(1) $\mathbb{R}^d$ and $\mathbb{C}^d$, as we learned in elementary real analysis.
(2) $C([a, b])$ with $\|f\| = \sup_{x\in[a,b]}|f(x)|$ is Banach, because a sequence $\{f_n\}$ is Cauchy iff it converges uniformly, and we know the uniform limit of continuous functions is continuous.

$C([a, b])$ with norm

$$\|f\|_{L^1} = \int_a^b |f(x)|\,\mathrm{d}x$$

is *not* complete, and therefore not Banach! This will verify the statement we made last lecture, that these spaces aren't the same. This is interesting behavior, because it doesn't happen in finite dimensions, and is an example of the subtle differences in behavior between finite-dimensional and infinite-dimensional vector spaces.

We'll let $a = -1$ and $b = 1$, though by suitable rescaling or translation this works for any $[a, b]$ with $a$ and $b$ finite.

[2] This was all that the professor said about the proof that the norm is continuous. Here's an alternate proof in case you, like me, didn't get it: since $x_n \to x$, then for any $n \in \mathbb{N}$, there's an $N_n$ such that if $m \ge N_n$, then $x_m - x \in B_{1/n}(0)$. But that means that $\|x_m - x\| < 1/n$. Since $1/n \to 0$, then $\|x_n - x\| \to 0$ as well.

Let $f_n(x)$ be 1 on $[-1,0]$, then decrease linearly on $[0,1/n]$, and then be 0 on $[1/n,1]$. Then,

$$\begin{aligned}\|f_n - f_m\|_{L^1} &= \int_{-1}^{1} |f_n(x) - f_m(x)|\,\mathrm{d}x \\ &= \int_0^1 |f_n(x)-f_m(x)|\,\mathrm{d}x \\ &\le \int_0^1 (|f_n(x)| + |f_m(x)|)\,\mathrm{d}x \\ &= \frac{1}{2n} + \frac{1}{2m}.\end{aligned}$$

This goes to 0, so $\{f_n\}$ is Cauchy. But it converges to the step function

$$f(x) = \begin{cases} 1, & x < 0 \\ 0, & x > 0. \end{cases}$$

This is because

$$\begin{aligned}\|f_n - f\|_{L^1} &= \int_{-1}^{1} |f_n(x) - f(x)|\,\mathrm{d}x \\ &= \int_0^1 |f_n(x)|\,\mathrm{d}x = \frac{1}{2n},\end{aligned}$$

which goes to 0, so $f_n \to f$ after all.

This means that when we talk about $C([a,b])$, unless otherwise specified, we'll use the other norm, which makes it into a Banach space.

This situation, where the same vector space has two norms with different topological properties, is actually fairly common.

**Definition.** Let $X$ be a vector space and $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on $X$. One says that the two norms are *equivalent* if there exist $c, d > 0$ such that for all $x \in X$, $c\|x\|_1 \le \|x\|_2 \le d\|x\|_1$.

This means that, though they might not agree precisely, the vague notions of "small" and "large" are the same in both norms.

We'll see eventually that all norms on a finite-dimensional space are equivalent, even though we already know that $\|\cdot\|$ and $\|\cdot\|_{L^1}$ are inequivalent on $C([a,b])$. We do know, however, that for $f \in C([0,1])$, $\|f\|_{L^1} \le \|f\|$,[3] but the other bound fails: there is no constant $C$ such that $\|f\| \le C\|f\|_{L^1}$. We'll see this using the sequence $\{f_n\}$, where $f_n$ increases linearly from 0 to $n$ on $[0,1/n]$, decreases on $[1/n,2/n]$, and is 0 elsewhere. This sweeps out a triangle, so $\|f_n\| = n$, but $\|f_n\|_{L^1} = 1$ for all $n$, and thus no such $C$ exists.

**Proposition 2.3.** *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two equivalent norms on $X$. Then, their induced topologies are the same.*

To be precise, the collections of open sets $\mathscr{O}_1$ and $\mathscr{O}_2$ induced from $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively, are identical.

*Proof.* We'll let $B_r^1(x)$ denote the ball of radius $r$ around $x$ in $\|\cdot\|_1$, and define $B_r^2(x)$ similarly.

Since $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, there exist $c$ and $d$ such that for any $x$ and $r$, $B^1_{r/d}(x) \subseteq B^2_r(x) \subseteq B^1_{r/c}(x)$. Thus, if $O_2$ is any open set in $\mathscr{O}_2$, then for any $x \in O_2$, there's an $r$ such that $B_r^2(x) \subseteq O_2$, and therefore $B^1_{r/d}(x) \subseteq O_2$, and so $O_2$ is open in $\mathscr{O}_1$, and the argument in the other direction is similar. ⊠

**Convexity.** Convexity is an important notion because it allows us to talk about the line joining two points.

**Definition.** Let $X$ be a vector space over $\mathbb{F}$. Then, a set $C \subseteq X$ is convex if whenever $x, y \in C$, the line $\{tx + (1-t)y : 0 \le t \le 1\}$ is contained in $C$.

**Proposition 2.4.** *In any NLS, $B_r(x)$ is convex.*

[3] More generally, on $C([a,b])$, $\|f\|_{L^1} \le (b-a)\|f\|$.

*Proof.* Let $y, z \in B_r(x)$ and $t \in [0,1]$. We want to show that $ty + (1-t)z \in B_r(x)$. We'll have to write $x$ as $x + tx - tx$ and then use the triangle inequality. Specifically,

$$\begin{aligned}\|ty + (1-t)z - x\| &= \|t(y-x) + (1-t)(z-x)\| \\ &\le t\|y-x\| + (1-t)\|z-x\| \\ &< tr + (1-t)r = r.\end{aligned}$$ ⊠

This is more interesting than it looks, because in some spaces that are otherwise similar to NLSes, there exist balls that are non-convex.

Even in finite dimensions, balls aren't necessarily round; they can even be square! But that doesn't make much of a difference.

**Linear Operators.** We'll talk about linear operators in order to manipulate and transform functions.

**Definition.** A *linear operator* is a function $T : X \to Y$ of vector spaces $X$ and $Y$ such that

(1) $T(x + y) = T(x) + T(y)$, and
(2) $T(\lambda x) = \lambda T(x)$.

The idea is that scalar multiplication and addition in $X$ and $Y$ (which are *a priori* very different) are considered the same by $T$, which commutes with them.

**Definition.** A linear operator $T : X \to Y$, where $X$ and $Y$ are NLSes, is *bounded* if it takes bounded sets to bounded sets.

That is, if $C \subseteq X$ is bounded, then $T(C) = \{y : y = T(x) \text{ for some } x \in C\}$.

The definition is nice, but everybody thinks of bounded operators by the following characterization.

**Proposition 2.5.** *Let $X$ and $Y$ be normed linear spaces and $T : X \to Y$ be linear. Then, $T$ is bounded iff there exists an $C > 0$ such that $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$.*

*Proof.* First, suppose $T$ is bounded. Then, the image of $B_1(0)$ (in $X$) is some bounded set, and therefore contained in a ball $B_R(0)$ for some $R$. In particular, if $y \in B_1(0)$, then $\|Ty\|_Y \le R$.

Given $x \in X$, if $x = 0$ then $Tx = 0$, so we're good. If $x \ne 0$, let $y = (1/2\|x\|_X) \cdot x$, so that $\|y\| = 1/2$, and therefore $y \in B_1(0)$, and therefore $\|Ty\| \le R$. That is,

$$\left\| T\left(\frac{1}{2\|x\|}\|x\|\right)\right\| = \frac{1}{2\|x\|}\|Tx\| \le R,$$

and therefore $\|Tx\| \le 2R\|x\|$, so with $C = 2R$ we're done.

Conversely, suppose there exists a $C > 0$ such that $\|Tx\| \le C\|x\|$ for all $x \in X$. Let $M \subseteq X$ be bounded; then, $M \subseteq B_R(0)$ for some $R$. For an $x \in M$, $\|Tx\| \le C\|x\| \le CR$, so $T(X) \subseteq B_{CR}(0)$ in $Y$, and thus $T$ is bounded. ⊠

Lecture 3.

## Bounded Linear Operators: 8/31/15

Let $X$ and $Y$ be normed linear spaces; the maps between them that we'll consider are linear operators $T : X \to Y$, as in the previous lecture.

If $T$ is one-to-one and onto, then we should have an inverse $T^{-1} : Y \to X$. It's easy to check that $T^{-1}$ is linear; you probably checked this as an undergraduate. In this situation, we have structure preservation: it doesn't matter whether you check addition in $X$ or in $Y$, or scalar multiplication. Thus, in the sense of linear algebra, $X$ and $Y$ look the same; they have the same addition and scalar multiplication. In this case, we say that $X$ and $Y$ are *isomorphic*; they may be unequal as sets (e.g. sequences or functions), but identical from the perspective of linear algebra.

For vector spaces, these maps are pretty cool, but for topology, we care about continuous maps $f : X \to Y$. Thus, as you might guess, when studying normed linear spaces, we care about maps $X \to Y$ that are both linear and continuous.

**Definition.** If $X$ and $Y$ are NLSes, then $B(X, Y)$ denotes the set of functions $f : X \to Y$ that are both linear and continuous.

Continuity means that for all $\varepsilon > 0$ there exists a $\delta > 0$ depending on $x$ and $\varepsilon$ such that when $d(x,y) < \delta$, then $d(f(x), f(y)) \le \varepsilon$. But since there's a norm defining the metric, this is equivalent to stating that when $\|x-y\| < \delta$, then $\|f(x) - f(y)\| \le \varepsilon$. And if $f = T$ is a linear operator, then $\|T(x) - T(y)\| < \varepsilon$ is equivalent to requiring $\|T(x-y)\| \le \varepsilon$. In other words, this doesn't depend on $x$ at all: letting $z = x - y$, continuity of a linear $T : X \to Y$ means that when $\|z\| < \delta$, then $\|Tz\| \le \varepsilon$.

In other words, if you know what a linear map does around 0, you know what it looks like everywhere.

**Proposition 3.1.** *Let $X$ and $Y$ be NLSes and $T : X \to Y$ be linear. Then, the following are equivalent:*

*(1) $T$ is continuous.*
*(2) $T$ is continuous at some $x_0 \in X$.*
*(3) $T$ is bounded.*

This is why we used the notation $B(X,Y)$: it stands for "bounded." And we can now talk about bounded linear maps, with continuity understood.

*Proof.* Clearly, (1) $\implies$ (2). For (2) $\implies$ (3), suppose $T$ is continuous at some $x_0 \in X$. With $\varepsilon = 1$, this means there's a $\delta > 0$ such that $\|x - x_0\| \le \delta$ implies $\|Tx - Tx_0\| \le 1$, i.e. $\|T(x - x_0)\| \le 1$. In other words, with $z = x - x_0$, when $\|z\| \le \delta$, we have $\|Tz\| \le 1$.

For $x = 0$ boundedness is clear, but if $x \ne 0$, then

$$\begin{aligned}\|Tx\|_Y &= \left\| \frac{\|x\|}{\delta} T\left(\frac{\delta x}{\|x\|}\right) \right\|_Y \\ &= \frac{\|x\|}{\delta} \left\| T\left(\frac{\delta x}{\|x\|}\right) \right\| \le \frac{1}{\delta}\|x\|_X,\end{aligned}$$

so with $C = 1/\delta$, $T$ is a bounded operator.

For (3) $\implies$ (1), we know $\|Tx\|_Y \le C\|x\|_X$ for some fixed $C$ and all $x \in X$. Let $\varepsilon > 0$ and pick any $x_0 \in X$. Then, if $\delta = \varepsilon / C$ and $\|x - x_0\| \le \delta$, then

$$\|T(x - x_0)\| \le C\|x - x_0\| \le C\delta = \varepsilon,$$

so $T$ is continuous at $x_0$ and therefore everywhere. ⊠

It turns out $B(X,Y)$ is a vector space itself, with $(f+g)(x) = f(x) + g(x)$ and $(\lambda \cdot f)(x) = \lambda \cdot (f(x))$, which is little surprise. But we do have to check that if $f = T$ and $g = S$ are linear, $f + g$ and $\lambda f$ are also linear, i.e. $(T+S)(x+y) = (T+S)(x) + (T+S)(y)$, and similarly for scalar multiplication.

What makes this more interesting is that $B(X,Y)$ is an NLS itself. What's the norm, you ask? Excellent question. The norm is

$$\|T\| = \|T\|_{B(X,Y)} = \sup_{x \in B_1(0)} \|Tx\|_Y.$$

Since $T$ is continuous and bounded, $T(B_1(0))$ is a bounded set. Then, the norm of $T$ is the radius of the smallest ball that contains $T(B_1(0))$, which is the supremum of the amount that $T$ scales any point in the unit ball. Since $T$ is bounded, the norm is a finite, nonnegative number.

Note that, even though we called this a norm, we still have to check that it's a norm!

**Proposition 3.2.** *Let $X$ and $Y$ be NLSes. Then, $\|\cdot\|_{B(X,Y)}$ is a norm on $B(X,Y)$. Moreover, if $T \in B(X,Y)$,*

$$\|T\| = \sup_{\|x\|_X \le 1} \|Tx\|_Y = \sup_{\|x\|_X = 1} \|Tx\|_Y = \sup_{x \ne 0} \frac{\|Tx\|_X}{\|x\|_X}.$$

*Furthermore, if $Y$ is Banach, then $B(X,Y)$ is too.*

This last point is quite interesting: completeness follows when the range is complete, but the domain doesn't matter.

*Proof.* First, that $\|\cdot\|$ is a norm: we have three properties to show.

- We need $\|T\| = 0$ iff $T = 0$. Clearly, if $T = 0$ (i.e. $T(x) = 0$ for all $x$), then $\|T\| = \sup_{x \in B_1(0)} \|Tx\| = \|0\| = 0$. Conversely, if we assume $\|T\| = 0$, then for any $x \in B_1(0)$, $\|Tx\| = 0$, so $Tx = 0$. Thus, $T|_{B_1(0)} = 0$. For general $x$, we'll scale $x = 2\|x\|(x/2\|x\|)$, so

$$Tx = 2\|x\| T\left(\frac{x}{2\|x\|}\right) = 2\|x\| \cdot 0 = 0,$$

since $x/2\|x\| \in B_1(0)$. Thus, $T = 0$.

- For linearity of the norm,

$$\|\lambda T\| = \sup_{x\in B_1(0)} \|\lambda T x\| = \sup_{x\in B_1(0)} |\lambda|\|T x\| = |\lambda| \sup_{x\in B_1(0)} \|T x\| = |\lambda|\|T\|.$$

**Exercise.** Finish the proof that this is a norm by addressing the triangle inequality, which isn't too complicated.

Next, we have the different ways of calculating the norm. The idea is that since $T$ is continuous, the supremum shouldn't depend on whether the boundary is present or not. One interesting corollary of the formulas for calculating $\|T\|$ is that for any $x \in X$, $\|T x\| \le \|T\|\|x\|$.

The last part does require care. Let $\{T_n\}_{n=1}^\infty$ be a Cauchy sequence. That is, given an $\varepsilon > 0$, there's an $N > 0$ such that if $m, n \ge N$, then $\|T_n - T_m\|_{B(X,Y)} \le \varepsilon$. Thus, given an $x \in X$, $\|T_n x - T_m x\|_Y \le \|T_n - T_m\|\|x\|_X$. The right-hand side goes to 0 as a Cauchy sequence in $m$ and $n$, and therefore the left-hand side does too. That is, $\{T_n x\}_{n-1}^\infty \subset Y$ is a Cauchy sequence. Since $Y$ is Banach, this means there's a limit $\lim_{n\to\infty} T_n x = T(x) \in Y$. This defines a map $T : X \to Y$; we need to prove that it's bounded linear and that $T_n \to T$.

First, let's look at linearity.

$$T(x + y) = \lim_{n\to\infty} T_n(x + y) = \lim_{n\to\infty} (T_n x + T_n y).$$

Since addition is continuous, we can break this up as

$$= \lim_{n\to\infty} T_n x + \lim_{n\to\infty} T_n y = T x + T y.$$

Similarly, since scalar multiplication is continuous,

$$T(\lambda x) = \lim_{n\to\infty} T_n(\lambda x) = \lambda T(x).$$

Next, let's check that $T$ is bounded. Since the norm is continuous,

$$\begin{aligned}\|T x\|_Y &= \left\|\lim_{n\to\infty} T_n x\right\|_Y \\ &= \lim_{n\to\infty} \|T_n x\|_Y.\end{aligned}$$

However, this limit *a priori* might not exist, so we have to use the lim sup.

$$\begin{aligned}&\le \limsup_{n\to\infty} \|T_n\|\|x\|_X \\ &= M\|x\|_X.\end{aligned}$$

Here, $M$ is an upper bound on $\|T_n\|$, because $\{T_n\}$ is Cauchy and therefore bounded. Thus, we know $T \in B(X, Y)$.

Finally, to show $T_n \to T$, we need to be careful: limits depend on the topology that we're using, and so we should be careful that we're using the topology defined by $\|\cdot\|_{B(X,Y)}$.

Let $x \in B_1(0)$. Then,

$$\begin{aligned}\|T x - T y\|_Y &= \lim_{m\to\infty} \|T_m x - T_n x\| \\ &= \lim_{m\to\infty} \|(T_m - T_n)x\| \\ &\le \limsup_{m\to\infty} \|T_m - T_n\|\|x\|.\end{aligned}$$

Since $\{T_n\}$ is Cauchy, then for any $\varepsilon > 0$, $\|T_m - T_n\| \le \varepsilon$ when $m, n$ are sufficiently large, and therefore the lim sup goes to 0 as $n \to \infty$, and so $T_n \to T$. ⊠

There's one particularly important case, in which $Y = \mathbb{F}$.

**Definition.** The *dual space* of an NLS $X$ is $X^* = B(X, \mathbb{F})$.

By Proposition 3.2, $X^*$ is always a Banach space.

Though $B(X, Y)$ can be complicated for general $Y$, one can often understand it more easily using $X^*$.

**Example 3.3.** We can connect this with finite-dimensional linear algebra that we're more familiar with, and see that it's actually quite special.

Let $X$ be a $d$-dimensional vector space over $\mathbb{F}$ with basis $\{e_n\}_{n=1}^d$. Thus,

$$\begin{aligned}X &= \operatorname{span}\{e_1, \dots, e_d\} \\ &= \{\alpha_1 e_1 + \cdots + \alpha_d e_d \mid \alpha_i \in \mathbb{F}\},\end{aligned}$$

and we can write $x = x_1e_1 + \cdots + x_de_d \in X$. The map $T : X \to \mathbb{F}^d$ sending $x \mapsto (x_1, \dots, x_d)$ is one-to-one, onto, and linear, so all finite-dimensional vector spaces over a specified field are isomorphic. Moreover, we showed that all norms over a finite-dimensional vector space are equivalent, so as NLSes, they're all isomorphic too! There are many norms, which may still be interesting, but there's only one topology.

Lecture 4.

## $\ell^p$-norms: 9/2/15

Recall that we were looking at examples of Banach spaces, and that the first examples we saw (Example 3.3) were finite-dimensional vector spaces. If $d = \dim X$ is finite, so that $X = \operatorname{span}\{e_1, \dots, e_n\}$ (which is a basis for $X$), then the map $T : X \to \mathbb{F}^d$ sending $(x_1e_1 + \cdots + x_de_d) \mapsto (x_1, \dots, x_d)$ is an isomorphism of vector spaces, and the claim is that these maps define the same topology as well.

But first, let's define some norms on $\mathbb{F}^d$. Let $1 \le p \le \infty$, and define

$$\|x\|_{\ell^p} = \begin{cases} \left(\displaystyle\sum_{n=1}^{d} |x_n|^p\right)^{1/p}, & p < \infty \\ \max_n |x_n|, & p = \infty. \end{cases}$$

Sometimes, these are denoted $\|x\|_{\ell_p}$. Also, the case $p = 2$ is our familiar Euclidean norm $\|x\|_{\ell^2} = |x|$.

We do have to show that these are norms. When $p = 1, \infty$, it's a straightforward check, and when $1 < p < \infty$, the first two properties are pretty simple, but the triangle inequality is harder.

**Lemma 4.1** (Young's inequality[4]). *Let $1 < p < \infty$ and $q$ be the* conjugate exponent *defined such that $1/p + 1/q = 1$. If $a, b \ge 0$, then $ab \le a^p/p + b^q/q$, with equality iff $a^p = b^q$. Moreover, for all $\varepsilon > 0$, there exists a $C$ depending on $p$ and $\varepsilon$ such that $ab \le \varepsilon a^p + Cb^q$.*

*Proof.* The proof is easy once you know the trick, to look at the right function. Let $u : [0, \infty) \to \mathbb{R}$ send

$$u(t) = \frac{t^p}{p} + \frac{1}{q} - t.$$

Its derivative is well-defined: $u'(t) = t^{p-1} - 1$, so $u'(0) = 1$. In particular, $u(0) = 1/q$, and $u(1) = 0$ is a strict minimum.

We'll apply this to $t = ab^{-q/p}$:

$$\begin{aligned} 0 \le u(ab^{-q/p}) &= \frac{a^p}{pb^q} + \frac{1}{q} - \frac{a}{b^{q/p}} \\ &= \frac{1}{b^q}\left(\frac{a^p}{p} + \frac{b^q}{q} - \frac{ab^q}{b^{q/p}}\right), \end{aligned}$$

but $b^q/b^{q/p} = b$, since $q - q/p = q(1 - 1/p) = 1$. Thus, $0 \le a^p/p + b^q/q - ab$, and equality holds iff $t = ab^{-q/p} = 1$, where $u(t)$ is equal to 0.

For the second part, we can write

$$ab = \big((\varepsilon p)^{1/p} a\big)\big((\varepsilon p)^{-1/p} b\big) \le \frac{\varepsilon p a^p}{p} + \frac{(\varepsilon p)^{-q/p}}{q} b^q. \qquad \boxtimes$$

For conjugate exponents, we have the convention that the conjugate of 1 is $\infty$, and vice versa.

**Theorem 4.2** (Hölder's inequality). *Let $1 \le p \le \infty$ and $q$ be its conjugate exponent. If $x, y \in \mathbb{F}^d$, then*

$$\sum_n |x_n y_n| \le \|x\|_{\ell^p} \|y\|_{\ell^q}.$$

When $p = 2$, this is also known as the *Cauchy-Schwarz inequality*.

[4] Young's inequality technically refers to a more general statement; this could be called "Young's inequality for products."

*Proof.* The cases $p = 1, \infty$ are trivial; expand their definitions out. Similarly, if $x = 0$ or $y = 0$, there's not a lot to say. Thus, we're left with $1 < p < \infty$, so we can use Lemma 4.1.

Let $a = |x_n|/\|x\|_{\ell^p}$ and $b = |y_n|/\|y\|_{\ell^q}$. Then, by Lemma 4.1,

$$\frac{|x_n|}{\|x\|_{\ell^p}}\frac{|y_n|}{\|y\|_{\ell^q}} \le \frac{|x_n|^p}{p\|x\|_{\ell^p}^p} + \frac{|y_n|^q}{q\|y\|_{\ell^q}^q},$$

so summing all $n$ of those,

$$\begin{aligned}\frac{\sum_n |x_n y_n|}{\|x\|_{\ell^p}\|y\|_{\ell^q}} &\le \frac{\sum |x_n|^p}{p\|x\|_{\ell^p}^p} + \frac{\sum |y_n|^q}{q\|y\|_{\ell^q}^q}\\ &= \frac{\|x\|_{\ell^p}^p}{p\|x\|_{\ell^p}^p} + \frac{\|y\|_{\ell^q}^q}{q\|x\|_{\ell^q}^q}\\ &= \frac{1}{p} + \frac{1}{q} = 1.\end{aligned}$$

⊠

Now, we can use this to prove the triangle inequality for $\|\cdot\|_{\ell^p}$. We'll need two things for the Hölder inequality, so just take one term out of the $p^{\text{th}}$ power:

$$\begin{aligned}\|x+y\|_{\ell^p}^p &= \sum_{n=1}^d |x_n + y_n|^p\\ &\le \sum_{n=1}^d |x_n + y_n|^{p-1}(|x_n| + |y_n|)\\ &\le \left(\sum_{n=1}^d |x_n + y_n|^{(p-1)q}\right)^{1/q}(\|x\|_{\ell^p} + \|y\|_{\ell^q}).\end{aligned}$$

Since $p$ and $q$ are conjugate, $p = (p-1)q$, so the first term is $\|x - y\|_{\ell^p}^{p/q}$. Thus,

$$\|x+y\|_{\ell^p}^{p-p/q} \le \|x\|_{\ell^p} + \|y\|_{\ell^p},$$

and $p - p/q = 1$, so we're done.

Moreover, all these norms are equivalent.

**Proposition 4.3.** *Let $1 \le p \le \infty$. Then, for all $x \in \mathbb{F}^d$,*

$$\|x\|_{\ell^\infty} \le \|x\|_{\ell^p} \le d^{1/p}\|x\|_{\ell^\infty}.$$

These estimates are sharp, the first at $x = (1, 0, 0 \dots, 0)$, and the second at $x = (1, 1, \dots, 1)$.

*Proof.* Let $m$ be an index for which $|x_m| = \max_n |x_n|$. Since $f(x) = x^{1/p}$ is an increasing function,

$$\|x\|_{\ell^\infty} = |x_m| = (|x_m|^p)^{1/p} \le \left(\sum_{n=1}^d |x_n|^p\right)^{1/p} = \|x\|_{\ell^p},$$

and

$$\begin{aligned}\|x\|_{\ell^p} &= \left(\sum_{n=1}^d |x_n|^p\right)^{1/p}\\ &\le \left(\sum_1^d |x_m|^p\right)^{1/p}\\ &= (d|x_m|^p)^{1/p} = d^{1/p}\|x\|_{\ell^\infty}.\end{aligned}$$

⊠

Notice that some of these proof methods fail horribly in infinite dimensions.

It turns out that on all finite-dimensional vector spaces, all norms are equivalent.

**Proposition 4.4.** *All norms on a finite-dimensional NLS are equivalent. Moreover, a $K \subset X$ is compact iff it is closed and bounded.*

That means there's only one topology.

*Proof.* Let $d = \dim X$ and $\{e_n\}_{n=1}^d$ be a basis. Then, let $T : X \to \mathbb{F}^d$ be the coordinate map defined above. Let $\cong$ denote an isomorphism of NLSes.

We'll define a norm $\|\cdot\|_1$ on $x$ by $\|x\|_1 = \|Tx\|_{\ell^1}$: of the three properties, the last two are trivial (since $T$ is linear), so we just need to prove that $\|x\|_1 = 0$ iff $x = 0$. But $T$ is one-to-one and onto, so this follows, and $\|\cdot\|_1$ is in fact a norm.[5]

Thus, $(X, \|\cdot\|_1) \cong (\mathbb{F}^d, \|\cdot\|_{\ell^1})$, so they really are the "same" space. This is because $T : X \to \mathbb{F}^d$ is a bounded map, with $C = 1$, and therefore continuous, and $T^{-1}$ is also linear and continuous. Thus, $T$ is an isomorphism of vector spaces and a homeomorphism of topological spaces, so we can take results in $\mathbb{F}^d$ and apply them to $X$.

The Heine-Borel theorem from undergraduate real analysis tells us that $K \subset \mathbb{F}^d$ is closed and bounded iff it's compact. But since $X$ and $\mathbb{F}^d$ have the same topology, then this is also true in $X$. In particular, $S_1^1 = \{x \in X : \|x\|_1 = 1\}$ is also compact.

Now, for any norm $\|\cdot\|$ on $X$ and $x \in X$,

$$\|x\| = \left\| \sum_{n=1}^d x_n e_n \right\| \le \sum_{n=1}^d |x_n| \|e_n\| \le C \|x\|_1,$$

where $C = \max_n \|e_n\|$. Notice that this step won't work in infinite dimensions. Our upper bound implies that $(Top)_{\|\cdot\|} \subseteq (Top)_{\|\cdot\|_1}$, so the former topology is said to be stronger. We'll prove the two are equal by providing a lower bound.

We have a continuous map $\|\cdot\| : (X, \|\cdot\|_1) \to \mathbb{R}$. It's also continuous as a map $\|\cdot\| : (X, \|\cdot\|) \to \mathbb{R}$. Let $a = \inf_{x \in S_1^1} \|x\|$; since $S^1$ is compact and the norm is continuous, the minimum is attained, and it must be positive (because $0 \notin S_1^1$).

Thus, for any $x \in X$, $\|x / \|x\|_1\| \ge a$, so $\|x\| \ge a\|x\|_1$, which is our desired lower bound. ⊠

**Corollary 4.5.** *If $X$ is a $d$-dimensional NLS, then $X \cong \mathbb{F}^d$.*

**Corollary 4.6.** *If $X$ and $Y$ are NLSes and $X$ is finite-dimensional, then every linear $T : X \to Y$ is bounded and $X^* = \mathbb{F}^d$, given by $T(x) = y \cdot x$.*

Lecture 5.

## $\ell^p$ and $L^p$-spaces: 9/4/15

*"There are different sizes of infinity, and this one is the best."*

Last time we showed that if $(X, \|\cdot\|)$ is a finite-dimensional NLS, then it's isomorphic and homeomorphic to $(\mathbb{F}^d, \|\cdot\|_{\ell^2})$, where $d = \dim X$. Moreover, $X$ is Banach, and $(\mathbb{F}^d)^* \cong \mathbb{F}^d$. Finite dimensions aren't very interesting, but they're a good place to gain intuition.

A lot of this nice stuff goes away for infinite-dimensional spaces, and some are nicer than others.

**Example 5.1.** Let $1 \le p \le \infty$. We'll define a space $\ell^p$ which behaves sort of like an "$\mathbb{F}^\infty$." Specifically,

$$\ell^p = \left\{ x = \{x_n\}_{n=1}^\infty : x_n \in \mathbb{F}, \|x\|_{\ell^p} < \infty \right\},$$

where

$$\|x\|_{\ell^p} = \begin{cases} \left( \displaystyle\sum_{n=1}^\infty |x_n|^p \right)^{1/p}, & p \text{ finite} \\ \sup_n |x_n|, & p = \infty. \end{cases}$$

The same proofs for the $\ell^p$-norms in finite-dimensional spaces apply, and show that $\ell^p$ is an NLS.

**Theorem 5.2** (Hölder's inequality in $\ell^p$). *If $1 \le p \le \infty$, $1/p + 1/q = 1$, and $x \in \ell^p$ and $y \in \ell^q$, then*

$$\sum_{n=1}^\infty |x_n y_n| \le \|x\|_{\ell^p} \|y\|_{\ell^q}.$$

[5] A great way to create a new norm is to map from one space to another (or the same one) and pull the norm back.

Again, the proof is identical to the one for the finite-dimensional $\ell^p$-norm.

Note that $\ell^\infty$ can be a bit weird relative to the rest of the $\ell^p$ spaces.

If $p$ is finite, then $\ell^p$ has countably infinite dimension, i.e. it has a basis that's countable. This is subtle: the span of a basis is the set of *finite* linear combinations; in the infinite case, we would have to worry about convergence. Anyways, set

$$e^{i_n} = \begin{cases} 1, & i = n \\ 0, & i \neq n. \end{cases}$$

Then, a basis for $\ell^p$, called the *Schauder basis*, is $\mathscr{B} = \{e^i\}_{i=1}^\infty$, and its span is

$$\operatorname{span}(\mathscr{B}) = \left\{\alpha_{i_1} e^{i_1} + \alpha_{i_2} e^{i_2} + \cdots + \alpha_{i_n} e^{i_n} : n \in \mathbb{N}, \alpha_{i_j} \in \mathbb{F}\right\}.$$

Note that this is *not* a basis in the linear-algebraic sense (which would have to be uncountable); rather, this means that $\ell^p$ is the closure of span$(\mathscr{B})$. That is, for all $x \in \ell^p$, there's a unique representation $x = \sum_{j=1}^\infty x_j e^j$, meaning that if $x_N$ denotes the $N^{\text{th}}$ partial sum, then $x_N \in \mathscr{B}$ for all $N$, and

$$\|x - x_N\|_{\ell^p} = \left(\sum_{n=N+1}^\infty |x_n|^p\right)^{1/p} \longrightarrow 0.$$

This is a little weird, but the point is that, since you can't take infinite sums in a basis, things can get a little strange. But everything comes from the finite case.

$\ell^\infty$ does *not* have a countable basis. As a result, we sometimes consider subspaces with a countable basis. Define

$$\begin{aligned} c_0 &= \{x \in \ell^\infty : \lim_{n\to\infty} x_n = 0\} \text{ and} \\ f_0 &= \{x \in \ell^\infty : x_n = 0 \text{ for all but finitely many } n\}. \end{aligned}$$

For example, $(1, 1, 1\dots) \in \ell^\infty$, but it's not in $c_0$ or $f_0$, and $(1, 1/2, 1/3, \dots)$ is in $c_0$ but not $f_0$. $f_0$ and $c_0$ inherit the $\ell^\infty$-norm and become NLSes in their own right.

If $1 \le p \le q < \infty$, then we have the following chain of inclusions:

$$f_0 \subseteq \ell^p \subseteq \ell^q \subseteq c_0 \subseteq \ell^\infty.$$

If you're looking for examples (or, sometimes, counterexamples), $c_0$ and $f_0$ are often useful. For example, on $f_0$, we have a function $T : f_0 \to \mathbb{F}$ defined by

$$T(x) = \sum_{n=1}^\infty n x_n.$$

Since each $\alpha \in f_0$ is a finite sequence, then this is well-defined, and it's linear, but it's not bounded, since $T(e^i) = i$ but $\|e^i\|_{\ell^\infty} = 1$ for all $i$. Thus, we have a linear map which is not continuous.

**Exercise.** If $1 \le p \le \infty$, show that $\ell^p$ is Banach.

This is conceptually easy but a bit of work, coming down to calculus, and so we know that limits of Cauchy sequences exist. However, since $\ell^1$ is a subspace of $\ell^\infty$, we can consider the NLS $(\ell^1, \|\cdot\|_{\ell^\infty})$; this space is not Banach.

**Lemma 5.3.** *Let $0 < p < 1$ and define $\ell^p$ in the same way as above. In this case, however, $\ell^p$ is not an NLS, because $\|\cdot\|_{\ell^p}$ isn't a norm.*

*Proof.* We can look at $(\mathbb{F}^2, \|\cdot\|_{\ell^p})$ to see this: we proved that, given the triangle inequality, the unit ball is convex. However, the unit ball isn't convex when $p < 1$. ⊠

The Hölder inequality allows us to create many continuous linear functionals $T : \ell^p \to \mathbb{F}$ when $1 \le p \le \infty$. Let $q$ be the conjugate exponent (so $1/p + 1/q = 1$), and choose any $y \in \ell^q$. Then, we can produce a $T_y \in (\ell^p)^*$, i.e. $T_y : \ell^p \to \mathbb{F}$, defined by

$$T_y(x) = \sum_{n=1}^\infty x_n y_n.$$

Moreover, $T_y$ is bounded, because $|T_y(x)| \le \|y\|_{\ell^q} \|x\|_{\ell^p}$.

This defines an inclusion $\ell^q \hookrightarrow (\ell^p)^*$.

**Exercise.** In fact, when $p$ is finite, $\ell^q = (\ell^p)^*$. Moreover, $T : \ell^q \to (\ell^p)^*$ sending $T(y) \to T_y$ is a bounded operator, as $\|T_y\|_{(\ell^p)^*} = \|y\|_{\ell^q}$.

That is, the dual space is the conjugate space; to show this, figure out how to write $T(e^i)$ as $y_i$ for some $y_i \in \ell^q$.
The above result is untrue for $\ell^\infty$; in fact, $(\ell^\infty)^* \supsetneqq \ell^1$, but $c_0^* = \ell^1$.
That's all that we really need to say about $\ell^p$ for now; it's one step up from finite-dimensional spaces, and is a bit different, but not all that exotic. Right now, our examples are $\mathbb{F}^d$, which is finite-dimensional; $\ell^p$ when $p$ is finite, which has countable dimension, and $\ell^\infty$, which has uncountable dimension.

**Lesbegue spaces.** Let $\Omega \subseteq \mathbb{R}^d$ be a measurable set. We want to define a space of functions on $\Omega$. However, when we talk about functions and measure, we really want to define two functions $f$ and $g$ as "the same" if $f(x) = g(x)$ except on a set of measure zero. If this is true, no integral can distinguish $f$ and $g$.

**Definition.** Let $1 \le p < \infty$, and define $L^p(\Omega)$ be the set of measurable functions[6] $f : \Omega \to \mathbb{F}$ such that $\int_\Omega |f(x)|^p \, \mathrm{d}x$ is finite. $L^p(\Omega)$ becomes an NLS with the norm

$$\|f\|_p = \left( \int_\Omega |f(x)|^p \right)^{1/p},$$

though we'll have to show that.

Once again, we can define this for $p < 1$, but it won't end up being a norm.
When $p = \infty$, we'll do things a little differently, as usual.

**Definition.**

- A measurable $f : \Omega \to \mathbb{F}$ is *essentially bounded* by $K \in \mathbb{R}$ if $|f(x)| \le K$ for almost every $x \in \Omega$ (i.e. the set where this is not true has measure zero).
- The *essential supremum* of $f$, denoted $\operatorname{ess\,sup}_{x \in \Omega} |f(x)|$, is the infimum of the $K$ that essentially bound $f$.

Then, we can define $L^\infty(\Omega)$ as the set of (equivalence classes of) measurable functions whose essential suprema are finite, and $\|f\|_\infty = \operatorname{ess\,sup}_{x\in\Omega} |f(x)|$. This will also be an NLS, though we'll have to show that too.

**Proposition 5.4.** *If $0 < p \le \infty$, then $L^p(\Omega)$ is a vector space, and $\|f\|_p = 0$ iff $f = 0$ almost everywhere on $\Omega$.*

*Proof.* First, why is $L^p(\Omega)$ closed under addition? If $p$ is finite, then

$$|f(x) + g(x)|^p \le (|f(x)| + |g(x)|)^p \le 2^p(|f(x)|^p + |g(x)|^p),$$

so when one integrates, if $f, g \in L^p(\Omega)$, then the rightmost quantity is bounded and therefore the leftmost one is. Scalar multiplication (and the scaling property of the norm) is easy: just write down the definition.
For $p = \infty$, the maximum of the sum cannot be bigger than the sum of the maxima, so $\|f+g\|_\infty = \|f\|_\infty + \|g\|_\infty$. Scaling and scalar multiplication are also straightforward. ⊠

Thus, all we have left is the triangle inequality, which we'll show next class.

Lecture 6.

## $L^p(\Omega)$ is Banach: 9/9/15

Recall that if $\Omega \subseteq \mathbb{R}^d$, then $L^p(\Omega)$ is the set of equivalence classes of measurable functions $\Omega \to \mathbb{F}$ with $\|f\|_p < \infty$, where $f \sim g$ if they differ on a set of measure zero. Then, the $p$-norm is

$$\|f\|_p = \begin{cases} \left( \int_\Omega |f(x)|^p \, \mathrm{d}x \right)^{1/p}, & p < \infty \\ \operatorname{ess\,sup}_{x\in\Omega} |f(x)|, & p = \infty. \end{cases}$$

Last time, we showed that $L^p(\Omega)$ is a vector space, and two of the properties of NLSes, the zero and scaling properties. Today we'll attack the triangle inequality; just as for $\ell^p$, we'll need Hölder's inequality.

**Proposition 6.1** (Hölder's inequality for $L^p$)**.** *Let $1 \le p \le \infty$ and $1/p + 1/q = 1$. If $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$ and $\|fg\|_1 \le \|f\|_p \|g\|_q$, with equality iff $|f(x)|^p$ is proportional to $|g(x)|^q$.*

[6] To be pedantic, the elements of $L^p(\Omega)$ are equivalence classes of functions that differ from $f$ on a set of measure zero, since the integrals are the same.

*Proof.* if $p = 1, \infty$, we already know that $\int_\Omega |f(x)g(x)|\,dx \le \|g\|_\infty \int_\Omega |f|\,dx = \|f\|_1\|g\|_\infty$.

If $1 < p < \infty$, we know from Lemma 4.1 that $ab \le a^p/p + b^q/q$, with equality when $a^p = b^q$. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then we're done; otherwise,

$$\frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_q} \le \frac{|f(x)|^p}{\|f\|_p^p p} + \frac{|g(x)|^q}{\|g\|_q^q q},$$

so integrating, we get

$$\frac{\int |fg|}{\|f\|_p\|g\|_q} \le 1,$$

with equality when $|f(x)|^p/\|f\|_p^p = |g(x)|^q/\|g\|_q^q$, which gives us our proportionality. ⊠

**Theorem 6.2** (Minkowski's inequality). *If $1 \le p \le \infty$, then $\|f+g\|_p \le \|f\|_p + \|g\|_p$.*

*Proof.* Notice that if $f$ or $g$ isn't in $L^p(\Omega)$, then its $p$-norm is infinite, so we're done. The result is also clear if $p = 1, \infty$: the supremum of the sum is less than the sum of the suprema, and similarly with absolute value.

So we only have to worry about $1 < p < \infty$, and here we'll use a similar trick as for $\ell^p$ spaces, taking one copy of a $p^{\text{th}}$ power.

$$\begin{aligned}
\|f+g\|_p^p &= \int_\Omega |f(x)+g(x)|^p\,dx \\
&\le \int_\Omega |f(x)+g(x)|^{p-1}(|f(x)|+|g(x)|)\,dx.
\end{aligned}$$

Using Hölder's inequality,

$$\begin{aligned}
&\le \left(\int_\Omega |f(x)+g(x)|^{(p-1)q}\right)^{1/q}\big(\|f\|_p + \|g\|_p\big) \\
&= \|f+g\|_p^{p-1}\big(\|f\|_p + \|g\|_p\big),
\end{aligned}$$

so dividing by $\|f+g\|^{p-1}$, we're done. ⊠

$L^p$ spaces are very important in analysis, and form an important set of examples for NLSes. A little later, we'll show that they're complete, but we should note that we're measuring the size of a function using varying $p$, which measure different things, between emphasizing large values at a point, or large values at infinity.

On $\mathbb{R}$, imagine a function that goes to $\infty$ as $x \to 0^+$ and 0 as $x \to \infty$. If $p$ is large, we're emphasizing the large values of the function, so if it grows too quickly it might not be in $L^p(\mathbb{R})$. If $p$ is small, then we're emphasizing the long tail as $x \to \infty$; if it dies too slowly, it might not be in $L^p(\mathbb{R})$. An instructive example is $x^p$, which is in some $L^q$ spaces but not others.

An easier way to think about this is to bound $\Omega$, so we don't have to worry about long tails.

**Proposition 6.3.** *Let $\mu$ denote the Lesbegue measure, and suppose $\mu(\Omega)$ is finite. Let $1 \le p \le q \le \infty$.*

*(1) If $f \in L^q(\Omega)$, then $f \in L^p(\Omega)$, and in fact $\|f\|_p = (\mu(\Omega))^{1/p-1/q}\|f\|_q$.*
*(2) If $f \in L^\infty(\Omega)$, then $f \in L^p(\Omega)$ for $1 \le p \le< \infty$, and $\lim_{p\to\infty}\|f\|_p = \|f\|_\infty$.*
*(3) If $f \in L^p(\Omega)$ for $1 \le p < \infty$ and $\|f\|_p \le K$ for all such $p$, then $f \in L^\infty(\Omega)$ and $\|f\|_\infty \le K$.*

These will be proven in the homework. Part (2) is the reason the $L^\infty$-norm is named such. Note also that there exist $f$ such that $f \in L^p(\Omega)$ for $1 \le p < \infty$ but $f \notin L^\infty(\Omega)$, even when $\Omega$ has finite measure.

The general proof idea is to consider sets of bad points and see what happens.

**Proposition 6.4.** *For $1 \le p \le \infty$ and $\Omega$ measurable, $L^p(\Omega)$ is complete.*

Thus, we have another useful class of Banach spaces.

*Proof.* As usual, we'll start with a Cauchy sequence $\{f_n(x)\}_{n=1}^\infty$ in $L^p(\Omega)$. The idea will be to write

$$f_n(x) = f_1(x) + f_2(x) - f_1(x) + f_3(x) - f_2(x) + \cdots + f_n(x) - f_{n-1}(x),$$

so if we group the $f_i(x) - f_{i-1}(x)$, then these pieces should be small, and therefore we ought to converge to some function $f(x)$. There are technical problems, though, since we don't know how fast the $f_n$ converge, so we need to try $f_i(x) - f_{i-k}(x)$ for $k > 1$. Moreover, we'll use absolute values. This is the idea; now, let's write it down carefully.

First, select a subsequence such that $\|f_{n_{j+1}} - f_{n_j}\| \le 2^{-j}$ for all $j$; we can do this because if wre have $n_{j-1}$, there's an $n_j$ such that $\|f_{n_j} - f_m\| \le 2^{-j}$ when $m \ge n_j \ge n_{j-1}$.

Let

$$F_m(x) = |f_{n_1}(x)| + \sum_{j=1}^{m} |f_{n_{j+1}}(x) - f_{n_j}(x)| \ge 0,$$

and additionally $\{F_m(x)\}$ is an increasing function, so there's a limit (which might be $\infty$, but that's OK). Let $F(x) = \lim_{m\to\infty} F_m(x) \in [0,\infty]$. Then,

$$\|F_m\|_p \le \|f_{n_1}\|_p + \sum_{j=1}^{n} 2^{-j} \le \|f_{n_1}\|_p + 1,$$

which is finite. But more interestingly, $F \in L^p(\Omega)$ too! We'll have to treat $L^\infty$ as a special case again.

If $p$ is finite, we'll use the monotone convergence theorem.

$$\begin{aligned}\int_\Omega |F(x)|^p \,\mathrm{d}x &= \int_\Omega \lim_{m\to\infty} |F_m(x)|^p \,\mathrm{d}x \\ &\le \lim_{m\to\infty} \int_\Omega |F_m(x)|^p \,\mathrm{d}x \\ &\le \|f_{n_1}\|_p + 1,\end{aligned}$$

which is finite.

When $p = \infty$, then $|F_m(x)| \le \|F_m\|_\infty \le \|f_{n_1}\|_\infty + 1$ for all $x \notin A_m$, where $\mu(A_m) = 0$. Thus, if $A = \bigcup_{n=1}^\infty A_n$, then $\mu(A) = 0$ too. Thus, $|F(x)| = \lim_{m\to\infty} |F_m(x)| \le K$ for some $K$ and all $m, x \notin A$, so $F \in L^\infty(\Omega)$.

Now,

$$f_{n_j+1}(x) = f_{n_1}(x) + (f_{n_2}(x) - f_{n_1}(x)) + \cdots + (f_{n_j+1}(x) - f_{n_j}(x)).$$

Thus, this converges absolutely *pointwise*[7] to some $f(x)$, so $f$ is measurable. Now, $|f_{n_j}(x)| \le F(x)$, so $|f(x)| \le F(x)$, and therefore $f \in L^p(\Omega)$.

But we need that $\|f_{n_j} - f\|_p \to 0$, so let's think about that. Again, we have to argue differently when $p = \infty$.

When $p$ is finite, we'll use the dominated convergence theorem on $|f_{n_j}(x) - f(x)| \le F(x) + |f(x)| \in L^p(\Omega)$:

$$\lim_{j\to\infty} \int_\Omega |f_{n_j}(x) - f(x)|^p \,\mathrm{d}x \le \int_\Omega \lim_{j\to\infty} |f_{n_j}(x) - f(x)|^p \,\mathrm{d}x \longrightarrow 0.$$

When $p$ is infinite, for any $j$ and $k$, there's a set $B_{n_j,n_k}$ with measure zero such that on $\Omega \setminus B_{n_j,n_k}$, $|f_{n_j}(x) - f_{n_k}(x)| \le \|f_{n_j} - f_{n_k}\|_\infty$. Thus,

$$B = \bigcup_j \bigcup_k B_{n_j,n_k}$$

is a countable union, so $\mu(B) = 0$. Since $\{f_{n_j}\}$ is Cauchy, then for any $x \notin B$ and $\varepsilon > 0$, there's an $N > 0$ such that if $j, k \ge N$, then $|f_{n_j}(x) - f_{n_k}(x)| \le \varepsilon$, so taking the pointwise limit $f_k(x) \to f(x)$, $|f_{n_j}(x) - f(x)| \le \varepsilon$. Thus, since we're avoiding $B$, $\|f_{n_j} - f\|_\infty \le \varepsilon$.

We're almost done: we have $f_{n_j} \to f$ in $L^p$, but we need $f_n \to f$ in $L^p$. If $\varepsilon > 0$, then there exists an $N > 0$ such that $\|f_n - f_{n_j}\|_p \le \varepsilon/2$ for all $n, n_j \ge N$. Therefore $|f_{n_j} - f|_p \le \varepsilon/2$ for all $n_j \ge N$, and therefore the triangle inequality tells us that

$$\|f_n - f\|_p \le \|f_n - f_{n_j}\|_p + \|f_{n_j} - f\|_p \le \varepsilon. \qquad \boxtimes$$

If you examine the proof, we've also proven an interesting result.

**Corollary 6.5.** *If $1 \le p < \infty$ and $\{f_n\}_{n=1}^\infty$ is a sequence in $L^p(\Omega)$ converging to $f$ in the $L^p$-norm, then there exists a subsequence $\{f_{n_j}\}_{j=1}^\infty$ such that $f_{n_j}(x) \to f(x)$ pointwise a.e.*

[7]We have multiple notions of convergence floating around; be careful to distinguish pointwise convergence, uniform convergence, and convergence in $L^p$.

So convergence in $L^p$ implies pointwise convergence of a subsequence almost everywhere. We'll use this later.

It turns out that the dual space to $L^p(\Omega)$ is $L^q(\Omega)$, where $q$ is the conjugate exponent. Given a $g \in L^q(\Omega)$, define an operator $T_g : L^p(\Omega) \to \mathbb{F}$ by

$$T_g(f) = \int_\Omega f(x)g(x)\,\mathrm{d}x,$$

which makes sense and is finite by Proposition 6.1. Thus, this is well-defined, and linear because the integral is. It's continuous, because it's bounded (by Hölder's inequality again): $T_g(f) \le \|g\|_q \|f\|_p$, so $\|t_g\| \le \|g\|_q$, and it's probably not a surprise that's actually an equality: choose something like $f(x) = |g(x)|^{q/p}/\|g\|_q$ (maybe with a power in the denominator), to see that the bound is sharp.

Thus, we've shown that $L^q(\Omega) \subseteq (L^p(\Omega))^*$ in some sense, for $1 \le p \le \infty$. However, if $p$ is finite, then $L^q(\Omega) = (L^p(\Omega))^*$; there are no other continuous linear functionals. When $p = \infty$, there are more, so the dual space is the space of positive measures: $g(x)\,\mathrm{d}x$ is a measure, but there are other measures that aren't of that form.

We won't prove this, but it follows from a deep theorem in analysis called the Radon-Nikodym theorem.

Lecture 7.

## The Hahn-Banach Theorem: 9/11/15

*"Almost everything has three properties. Have you noticed that?"*

**Corollary 7.1.** *Let $X$ be an NLS and $Y \subset X$ be a linear subspace. Then, there exists an $F \in X^*$ such that $F|_Y = f$ and $\|F\|_{X^*} = \|f\|_{Y^*}$.*

Though $L^p$ functions can be complicated, all of them can be well-approximated by less complicated functions. Recall that a *simple* function is a Lesbegue-integrable function that takes on only finitely many values, and that a function is *compactly supported* if it is equal to 0 outside of a compact set.

**Proposition 7.2.** *For $1 \le p \le \infty$, the set $\mathscr{S}$ of all measurable simple functions with compact support is dense in $L^p(\Omega)$.*

This says that for any $f \in L^p(\Omega)$ and $\varepsilon > 0$, there's a $\varphi \in \mathscr{S}$ such that $\|f - \varphi\|_{L^p(\Omega)} < \varepsilon$. The proof comes from measure theory: the integral was defined by the limit of approximations by simple functions, and so the integrals are successively better approximations.

**Definition.** Let $C_0^\infty(\Omega)$ denote the space of compactly supported, continuous functions.

**Proposition 7.3.** *If $\Omega$ is an open set and $1 \le p < \infty$, then $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$.*

The proof follows from another measure-theoretic result called Lusin's theorem.

Now, we'll move into some deeper (and, well, harder) theorems and questions in functional analysis. We'll start with a question.

Let $X$ be a finite-dimensional NLS and $Y \subset X$ be a subspace. Given a linear $f : Y \to \mathbb{R}$, can we extend $f$ to $X$? The answer is yes. But what about the infinite-dimensional case? Here, we care about continuous (so bounded) linear operators.

Once again, the answer is that it's possible, but this is hard to prove, and it'll take us a while to prove that. We won't need all the properties of a norm to prove that, so we can weaken what we need in terms of the norm.

**Definition.** Let $X$ be a vector space over $\mathbb{F}$. We say that $p : X \to [0, \infty)$ is *sublinear* if

(1) $p(\lambda x) = \lambda p(x)$ for all $\lambda \ge 0$ and $x \in X$, and
(2) $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$.

If in addition $p$ satisfies (1) for all $\lambda \in \mathbb{F}$, $p$ is called a *seminorm*.

If a seminorm also satisfies $p(x) = 0$ implies $x = 0$, then $p$ is a norm.

The Hahn-Banach theorem about extension of linear operators will apply perfectly well to sublinear operators. First, let's deal with the simplest version we can think of.

**Lemma 7.4.** *Let $X$ be a vector space over $\mathbb{R}$ and $Y \subsetneq X$ be a linear subspace. Let $p$ be sublinear on $X$ and $f : Y \to \mathbb{R}$ be linear such that $f(y) \le p(y)$ for all $y \in Y$. For a given $x_0 \in X \setminus Y$, let $\widetilde{Y} = \operatorname{span}\{Y, x_0\} = Y + \mathbb{R}x_0 = \{y + \lambda x_0 : y \in Y, \lambda \in \mathbb{R}\}$; then, there exists a linear map $\widetilde{f} : \widetilde{Y} \to \mathbb{R}$ such that $\widetilde{f}|_Y = f$ and $-p(-x) \le \widetilde{f}(x) \le p(x)$ for all $x \in \widetilde{Y}$.*

The definitions of $\widetilde{Y}$ all show that it's "$Y$ plus one more dimension."

*Proof.* If $\widetilde{f}(x) \le p(x)$, then $-\widetilde{f}(x) = \widetilde{f}(-x) \le p(-x)$, so $\widetilde{f}(x) \ge -p(-x)$, and so the lower bound comes for free.

We'll present the proof not as a cleaned-up proof, but how one would think of the proof when trying to prove it.

If we had such an $\widetilde{f}$, what would it look like? $\widetilde{y} \in \widetilde{Y}$ can be written $\widetilde{y} = y + \lambda x_0$ for some $y \in Y$ and $\lambda \in \mathbb{R}$, so $\widetilde{f}(\widetilde{y}) = \widetilde{f}(y + \lambda x_0) = \widetilde{f}(y) + \lambda \widetilde{f}(x_0) = f(y) + \lambda \widetilde{f}(x_0)$, since $\widetilde{f}|_Y = f$.

So if we had defined $\alpha \in \mathbb{R}$ to be $\widetilde{f}(x_0)$, then we get a function, and correspondingly, given $\widetilde{f}$, we get $\alpha = \widetilde{f}(x_0)$. Thus, $\widetilde{f}$ is characterized by $\alpha$.

However, we need to be careful: is this really well-defined? We chose $y$; what if you choose a different one than I do? It turns out that you have to choose the same $y$: suppose $\widetilde{y} = y + \lambda x_0 = z + \mu x_0$ for $y, z \in Y$ and $\lambda, \mu \in \mathbb{R}$. Thus, $y - z = (\mu - \lambda)x_0$, but $y - z \in Y$, so since $x_0 \notin Y$, then $\mu - \lambda = 0$, and therefore $y = z$; thus, this choice of $y$ is well-defined, so $\widetilde{f}$ really is characterized by $\alpha$.

So now we need to find an $\alpha$ such that $\widetilde{f}(\widetilde{y}) = f(y) + \lambda\alpha \le p(y + \lambda x_0)$. If $\lambda = 0$ this works, so let's focus on $\lambda \ne 0$. Rescale: let $y = -\lambda x$, so we want to show that $f(-\lambda x) + \lambda \cdot \alpha \le p(\lambda(x_0 - x))$, or $\lambda(-f(x) + \alpha) \le p(-\lambda(x - x_0))$.

If $\lambda < 0$, then divide by $-\lambda$: $f(x) - \alpha \le p(x - x_0)$; when $\lambda > 0$, we get a change in sign: $-(f(x) - \alpha) \le p(-(x - x_0))$. Together, this means $-p(-(x - x_0)) \le f(x) - \alpha \le p(x - x_0)$. Rearranging,

$$f(x) - p(x - x_0) \le \alpha \le f(x) + p(x_0 - x).$$

This is our requirement; that is, if there's an $\alpha$ that satisfies this for all $x \in Y$, then we have our desired linear functional.

So let $a = \sup_{x \in Y}(f(x) - p(x - x_0))$ and $b = \inf_{x \in Y} f(x) + p(x_0 - x)$. Now we can ignore $\alpha$ and ask, is it true that $a \le b$? If so, we're done.

Let $x, y \in Y$. Since $p$ is sublinear, then

$$\begin{aligned} f(x) - f(y) = f(x - y) &\le p(x - y) \\ &\le p(x - x_0) + p(x_0 - y) \\ \implies f(x) - p(x - x_0) &\le f(y) + p(x_0 - y). \end{aligned}$$

In the last equation, first take the infimum on the left, which is $a$, and the right side doesn't change; then, take the supremum on the right, which is $b$, and the left side doesn't change. Thus $a \le b$. ⊠

This proof can be shortened, by starting with $\alpha$ and suddenly magical things happen, but this helps it make more sense and feel more rigorous.

**Transfinite Induction and Generalizing Lemma 7.4.** Applying this inductively, we can extend a finite number of dimensions, and even a countable number of dimensions! However, standard induction doesn't allow us to extend by an uncountable number of dimensions. This will require a technique called transfinite induction, and therefore a brief vacation into set theory.

**Definition.** A *ordering* on a set $\mathscr{S}$ is a binary relation $\preceq$ such that for all $x, y, z \in \mathscr{S}$,

(1) $x \preceq x$,
(2) if $x \preceq y$ and $y \preceq x$, then $x = y$, and
(3) if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

Not every set can be ordered. However, some can be partially ordered; a *partial order* on a set is the same except that only some pairs $x \preceq y$ are defined, but the same order axioms are satisfied (in particular, $x \preceq x$ is always defined and true, and if $x \preceq y$ and $y \preceq z$, then $x \preceq z$ is defined and true). A *chain* in a partially ordered set $\mathscr{S}$ is a $\mathscr{C} \subset \mathscr{S}$ such that $\preceq|_{\mathscr{C}}$ is a total order: every element can be compared.

**Example 7.5.** On $\mathbb{C}$, write $z = r_z e^{i\theta_z}$, with $\theta_z \in [0, 2\pi)$.

(1) An ordering on $\mathbb{C}$ can be given by $x \preceq y$ iff $r_x < r_y$ or $r_x = r_y$ and $\theta_x \le \theta_y$.
(2) A partial ordering on $\mathbb{C}$ can be given by $x \preceq y$ iff $\theta_x = \theta_y$ and $r_x \le r_y$ (and is undefined if $\theta_x \ne \theta_y$).

We'll need a more complicated order, which requires using Zorn's lemma. This comes from an axiom of set theory called the Axiom of Choice, which states that, given any collection of nonempty sets, it's possible to choose one element out of each set.

Zorn's lemma is equivalent to the Axiom of Choice, but it somehow seems harder to believe.

**Lemma 7.6** (Zorn's lemma)**.** *Let $\mathscr{S}$ be a nonempty, partially ordered set, and suppose every chain $\mathscr{C} \subseteq \mathscr{S}$ has an upper bound, i.e. for all $\mathscr{C}$, there's a $u \in \mathscr{C}$ such that $x \preceq U$ for all $x \in \mathscr{C}$. Then, $\mathscr{S}$ has at least one maximal element $m$, i.e. if $m \preceq x$ for some $x \in \mathscr{S}$, then $x = m$.*

Next time, we'll use this to extend by an uncountable number of dimensions; then, we'll remove the requirement that the base field is real.

Lecture 8.

## The Hahn-Banach Theorem, II: 9/14/15

Recall that we're in the middle of proving the Hahn-Banach theorem, and therefore should remember the results we're going to need. We defined orders and partial orders and chains within partially ordered sets last lecture, and cited Zorn's lemma, Lemma 7.6, which gives conditions for when a partially ordered set has a maximal element. Finally, we have Corollary 7.1 in mind as a long-term goal.

Since we have a possibly countable number of dimensions, we have to use transfinite induction to prove the most general theorem, which is why Zorn's lemma shows up.

**Theorem 8.1** (Hahn-Banach theorem for real vector spaces)**.** *Let $X$ be a vector space over $\mathbb{R}$, $Y \subset X$ be a subspace, and $p$ be sublinear on $X$. If $f : Y \to \mathbb{R}$ is linear on $Y$ and $f(x) \le p(x)$ for all $x \in Y$, then there exists a linear $F : X \to \mathbb{R}$ such that $F|_Y = f$ and $-p(-x) \le F(x) \le p(x)$ for all $x \in X$.*

*Proof.* Let $\mathscr{S}$ denote the set of all linear extensions $g$ of $f$ to a subspace $D(g) \subset X$ containing $Y$, and such that $g(x) \le p(x)$ for all $x \in D(g)$. Since $f \in \mathscr{S}$, then $f$ is nonempty. We'll turn $\mathscr{S}$ into a partially ordered set by saying that $g \preceq h$ if $h$ extends $g$, i.e. $D(g) \subseteq D(h)$ and $h|_{D(g)} = g$.

Let $\mathscr{C}$ be a chain in $\mathscr{S}$, and let

$$D = \bigcup_{g \in \mathscr{C}} D(g).$$

Since these $D(g)$ are nested (i.e. one of $D(g) \subset D(h)$ or $D(g) \supset D(h)$ for all $g, h \in \mathscr{C}$), then $D$ is a vector space.[8] Then, we'll define $g_{\mathscr{C}}$ as follows: if $x \in D$, then $x \in D(g)$ for some $g \in \mathscr{C}$, so define $g_{\mathscr{C}}(x) = g(x)$. Is this well-defined? Yes, because if $x \in D(g) \cap D(h)$, then without loss of generality $g \preceq h$, and so $g(x) = h(x)$. Thus, we get a function $g_{\mathscr{C}} : D \to \mathbb{R}$, which is linear (which follows from its definition), and is bounded by $p$ (specifically, $g(x) \le p(x)$ for all $x \in D$), since each $g \in \mathscr{C}$ is. Thus, $g_{\mathscr{C}} \in \mathscr{C}$, and it's an upper bound for $\mathscr{C}$.

Applying Zorn's lemma, we have a maximal element $F$ for $\mathscr{S}$; since $F \in \mathscr{S}$, then it's a linear extension of $f$ and is bounded by $p$. So the final question is, what's $D(F)$? Suppose $D(F) \subsetneq F$; then, there exists some $x_0 \in X \setminus D(F)$, so by Lemma 7.4 we can extend $F$ to $\operatorname{span}\{D(F), x_0\}$. But this contradicts the fact that $D(F)$ is maximal. Thus, $D(F) = X$. ⊠

Awesome. Now, let's deal with complex vector spaces. Since we want scalar multiplication for all $\lambda \in \mathbb{C}$, we'll have to use a seminorm instead.

**Theorem 8.2** (Hahn-Banach theorem for complex vector spaces)**.** *Let $X$ be a vector space over $\mathbb{F}$, $Y \subset X$ be a linear subspace, and $p$ be a seminorm. If $f : Y \to \mathbb{F}$ is a linear functional such that $|f(x)| \le p(x)$ for all $x \in Y$, then there exists an extension $F : X \to \mathbb{F}$ such that $F|_Y = f$ and $|F(x)| \le p(x)$ for all $x \in X$.*

*Proof.* We'll assume $\mathbb{F} = \mathbb{C}$, since the real case comes from Theorem 8.1. Then, we can write $f(x) = g(x) + ih(x)$ for $g, h$ real linear, since

$$\begin{aligned} f(x+g) &= g(x+y) + ih(x+y) \\ &= f(x) + f(y) = g(x) + g(y) + ih(x) + ih(y), \end{aligned}$$

and scalar multiplication is similar, though only for real scalars. Instead, $f(ix) = if(x) = -h(x) + ig(x)$, and this is also $g(ix) + ih(ix)$. Thus, $h(x) = -g(ix)$. That is, since $f$ is linear, $f(x) = g(x) - ig(ix)$, which is a general fact.

Since $g$ is real linear, then Theorem 8.1 yields a real extension $G$ on $X$, because $|g(x)| \le |f(x)| \le p(x)$, and we have that $|G(x)| \le p(x)$.

[8] This is an important point; the union of subspaces isn't in general a vector subspace when they're not nested.

Define $F(x) = G(x) - iG(ix)$, which is a function $F : X \to \mathbb{C}$ that commutes with addition and real scalar multiplication. Thus, we need to check complex scalar multiplication, and therefore that $F(ix) = iF(x)$. Let's check that:

$$F(ix) = G(ix) - iG(-x) = G(ix) + iG(x) = i(G(x) - iG(ix)).$$

Therefore $F$ is $\mathbb{C}$-linear. Moreover, if $x \in Y$, then $F(x) = G(x) - iG(ix) = g(x) - ig(ix)$, and therefore $F|_Y = f$ as desired. Thus, the only thing left to check is the bound.

Let $x \in X$, and write $F(x) = re^{i\theta}$. Then,

$$r = |F(x)| = e^{-i\theta}F(x) = F\big(e^{-i\theta}x\big) = G\big(g^{-i\theta}(x)\big) - iG\big(-e^{-i\theta}x\big),$$

but the second term is imaginary, and therefore must be zero. Then,

$$\le p\big(e^{-i\theta}(x)\big) = |e^{-i\theta}|p(x) = p(x). \qquad \boxtimes$$

As a corollary, notice that $p(x) = \|f\|_{Y^*}\|x\|_X$.

The Hahn-Banach theorem has a great number of corollaries, which provide a lot of insight into NLSes and Banach spaces.

**Corollary 8.3.** *Let $X$ be an NLS and $x_0 \in X \setminus 0$ be fixed. Then, there exists an $f \in X^*$ such that $\|f\|_{X^*} = 1$ and $f(x_0) = \|x_0\|$.*

The idea is to define $f$ on a subspace where it's easy to define, and then extend.

*Proof.* Let $Z = \mathbb{F}x_0$, and define $h : Z \to \mathbb{F}$ by $h(\lambda x_0) = \lambda\|x_0\|$. Then, $|h(x_0)| = |\lambda|\|x_0\| = \|\lambda x_0\|$, so $|h(x)| \le \|x\|$ for all $x \in Z$ and $\|h\| = 1$. Then, we use Theorem 8.2 to extend $h$ to the desired $f$. $\boxtimes$

**Corollary 8.4.** *For any $\alpha \in \mathbb{F}$, there exists an $f \in X^*$ such that $f(x_0) = \alpha\|x_0\|$ (and therefore $\|f\|_{X^*} = \alpha$).*

The proof is the same as for Corollary 8.3, but one defines $h(\lambda x_0) = \alpha\lambda\|x_0\|$ instead.

Here's a more interesting corllary.

**Corollary 8.5.** *Let $X$ be an NLS and $x \in X$. Then,*

$$\|x\| = \sup_{\substack{f \in X^* \\ f \ne 0}} \frac{|f(x)|}{\|f\|_{X^*}} = \sup_{\substack{f \in X^* \\ \|f\|_{X^*}=1}} \frac{|f(x)|}{\|f\|_{X^*}}.$$

Often, when one knows the structure of the dual space better than that of the original space, this can be a useful way to calculate a norm.

*Proof.* For all $f \in X^*$ with $f \ne 0$, we know $|f(x)| \le \|f\|_{X^*} \le \|x\|$, so we know the supremum is still bounded by $\|x\|$. To get the other bound, we need the Hahn-Banach theorem, which says that there exists a $\widetilde{f} \in X^*$ such that $\widetilde{f}(x) = \|\widetilde{f}\|\|x\|$; then,

$$\sup_{\substack{f \in X^* \\ f \ne 0}} \frac{|f(x)|}{\|f\|_{X^*}} \ge \frac{\widetilde{f}(x)}{\|f\|} = \|x\|. \qquad \boxtimes$$

The idea here is that we can look at $\|x\|$, which is a calculation involving an abstract vector, or $\{|f(x)|\}_{f \in X^*}$, which is a collection of numbers, which sometimes is nicer. This is a common theme in functional analysis. The following result is related, at least in ideas.

**Proposition 8.6.** *If $f(x) = f(y)$ for all $f \in X^*$, then $x = y$.*

We'll prove this next time.

Lecture 9.

## Separability: 9/16/15

Recall that we're in the middle of exploring the consequences of the Hahn-Banach theorem, Theorems 8.1 and 8.2. For example, if $X$ is an NLS and $x_0 \in X$, then there's an $f \in X^*$ such that $f(x_0) = \|x_0\|$ (Corollary 8.4), that you can calculate $\|x\|$ from the norms of $f \in X^*$ (Corollary 8.5), and more.

**Proposition 9.1.** *If $X$ is an NLS, then $X^*$* separates *points in $X$, i.e. for any $x, y \in X$, there exists an $f \in X^*$ such that $f(x) \ne f(y)$, and if $f(x) = f(y)$ for all $f \in X^*$, then $x = y$.*

The recurring theme is that if you know what all the linear functionals do to an element, you know what that element is.

*Proof.* Choose $x, y \in X^*$ such that $x \neq y$. Then, $x - y = x_0 \in X$ and $x_0 \neq 0$, and there exists an $f \in X^*$ such that $f(x_0) \neq 0$, so $0 \neq f(x-y) = f(x) - f(y)$. ⊠

**Corollary 9.2.** *If $f(x) = 0$ for all $x \in X^*$, then $x = 0$.*

Oftentimes, one creates simple functionals by doing something interesting on a finite-dimensional subspace and then extending à la the Hahn-Banach theorem.

**Definition.** In an NLS $X$, the *distance* between a subspace $Y \subset X$ and a $w \in X$ is $\operatorname{dist}(w, Y) = \inf_{y \in Y} \|w - y\|$.

This is nonnegative, and sometimes it's zero even when $w \notin Y$.

**Lemma 9.3** (Mazur Separation Thm. I)**.** *Let $X$ be an NLS, $Y \subset X$ be a subspace, and $w \in X \setminus Y$. Suppose $d = \operatorname{dist}(w, Y) > 0$. Then, there exists an $f \in X^*$ with*

- *$\|f\| \leq 1$,*
- *$f(w) = d$, and*
- *$f(y) = 0$ for all $y \in Y$.*

*Proof.* Let $Z = Y + \mathbb{F}w$. Then, any $z \in Z$ has a unique representation as $z = y + \lambda w$ for exactly one choice of $y \in Y$ and $\lambda \in \mathbb{F}$ (which we discussed last time).

Then, define $g : Z \to \mathbb{F}$ by $g(y + \lambda w) = \lambda d$. $g$ is clearly linear, but it's less clear why $\|g\| \leq 1$.

$$\left| g\left(\frac{y + \lambda w}{\|y + \lambda w\|}\right)\right| = \frac{|\lambda| d}{\|y + \lambda w\|} = \frac{d}{\|(1/\lambda)y + w\|}.$$

Since $(1/\lambda)y \in Y$, then $\|(1/\lambda)y + w\| \geq d$, and therefore $d/\|(1/\lambda)y + w\| \leq 1$. Then, we use the Hahn-Banach theorem to extend to $X$. ⊠

We'll introduce another notion, entirely topological, which will be useful.

**Definition.** A topological space $X$ is *separable* if it contains a countable dense subset, i.e. a $\mathscr{D} \subset X$ such that $\overline{\mathscr{D}} = X$.

A space might be large and scary, but if it's separable, then everything is close, and therefore we can get a little control on it.

**Example 9.4.**

(1) $\mathbb{Q} \subset \mathbb{R}$. $\mathbb{Q}$ is countable and every real number can be arbitrarily well approximated by rational numbers, so $\mathbb{R}$ is separable.
(2) $\mathbb{Q}(i) = \mathbb{Q} + i\mathbb{Q} \subseteq \mathbb{C}$ is countable and dense, so $\mathbb{C}$ is separable.
(3) $\mathbb{F}^d$ is also separable, with the countable dense subset either $\mathbb{Q}^d$ or $\mathbb{Q}(i)^d$.
(4) If $1 \leq p < \infty$, our Schauder basis for $\ell^p$ is uncountable, but we can take instead the $\mathbb{Q}(i)$-span (or the $\mathbb{Q}$-span if $\mathbb{F} = \mathbb{R}$) of $\{e_i\}_{i=1}^{\infty}$; this is a countable dense subset of $\ell^p$, so $\ell^p$ is separable.
(5) If $1 \leq p < \infty$, then $L^p(\Omega)$ is separable. This one is a little more surprising. The set $S$ of simple functions (functions which are constant on a finite number of intervals). $S \subset L^p(\Omega)$ is dense, but uncountable, so we have to restrict it in two ways: first, restrict the allowed intervals to have rational coefficients, and then restrict the functions to take on values in $\mathbb{Q}(i)$ (or $\mathbb{Q}$; we'll assume that when we talk about $\mathbb{Q}(i)$, then we mean $\mathbb{Q}$ for $\mathbb{R}$). Thus restricted, we have our countable dense subset.

This argument doesn't work for $L^\infty(\Omega)$, since simple functions aren't dense in it, and in fact $L^\infty(\Omega)$ isn't separable.

**Proposition 9.5.** *Let $X$ be an NLS. If $X^*$ is separable, then $X$ is separable.*

The converse isn't true, because $L^1(\Omega)$ is separable but $L^\infty(\Omega)$ isn't. So if you start with a separable space, your dual might be bigger.

*Proof.* Let $\{f_n\}_{n=1}^{\infty}$ be a countable, dense subset of $X^*$. We'll use this to construct a countable, dense subset of $X$. Since $\|f\| = \sup_{\|x\|=1} |f(x)|$, then we can choose for each $n$ an $x_n$ such that $\|x_n\| = 1$ and $|f_n(x_n)| \geq (1/2)\|f_n\|$, giving us a sequence $\{x_n\}_{n=1}^{\infty}$.

Then, let $\mathscr{D} = \operatorname{span}_{\mathbb{Q}(i)}\{x_n\}$, which is still countable, and we'll show that $\overline{\mathscr{D}} = X$. Suppose that it weren't: then, there exists a $w \in X \setminus \overline{\mathscr{D}}$. Let $d = \operatorname{dist}(w, \overline{\mathscr{D}}) = \inf_{x\in\overline{\mathscr{D}}} \|w - x\| > 0$. If we can show that $\|w - y_n\| \to 0$ for some sequence $\{y_n\}_{n=1}^\infty$, then since $\overline{\mathscr{D}}$ is closed, that would imply $w \in \mathscr{D}$.

Since $\overline{\mathbb{Q}(i)} = \mathbb{C}$, then $\overline{\mathscr{D}}$ is a linear subspace of $X$; thus, by Lemma 9.3, there exists an $f \in X^*$ such that $f|_{\overline{\mathscr{D}}} = 0$ and $f(w) = d > 0$. But there's a sequence $\{f_{n_k}\}_{k=1}^\infty$ such that $f_{n_k} \to f$. Thus,

$$\|f_{n_k} - f\| \ge |f(x_{n_k}) - f_{n_k}(x_{n_k})| = |f_{n_k}(x_{n_k})| \ge \frac{1}{2}\|f_{n_k}\|.$$

Since $f_{n_k} - f \to 0$, then this means $f_{n_k} \to 0$, and so $f = 0$. But this is a contradiction. ⊠

So far, we've always looked at sets that are subspaces. Here's an example where we don't do that.

**Definition.** Let $X$ be an NLS and $C \subseteq X$ be a sub*set* (not necessarily a subspace). Then, $C$ is *balanced* if for any $\lambda \in \mathbb{F}$ with $|\lambda| \le 1$, then when $x \in C$, $\lambda x \in C$.

For example, if $\mathbb{F} = \mathbb{C}$, then this implies that $C$ is invariant under rotation, as well as contractions. Note that all subspaces are balanced.

**Lemma 9.6** (Mazur Separation Thm. II). *Let $X$ be an NLS and $C \subseteq X$ be a closed, convex, and balanced set. Then, for any $w \in X \setminus C$, there exists an $f \in X^*$ such that $|f(x)| \le 1$ for $x \in C$ and $f(w) > 1$.*

*Proof.* Since $C$ is closed and $w \notin C$, we can choose a ball $B + w$ about $w$ (so $B$ is a ball centered at the origin) such that $B \cap C = \emptyset$. Then, we can define the *Minkowski functional* $p : X \to [0, \infty)$ by

$$p(x) = \inf\left\{t > 0 : \frac{x}{t} \in C + B\right\}.$$

Here, $C + B$ is a slight fattening of our set $C$, but we can guarantee that $w \notin C + B$. Moreover, $0 \in C$, because $C$ is balanced; therefore, $p(x)$ is always finite. We also know that $p(x) \le 1$ if $x \in C$ and $p(w) > 1$.

Moreover, $p$ is a seminorm: since $C$ is balaced, $p(\lambda x) = p(|\lambda| x) = |\lambda| p(x)$. We also have the triangle inequality, which is left to the reader.

Now, we use Theorem 8.2: let $Y = \mathbb{F}w$, and if $f(\lambda w) = \lambda p(w)$, then $f(w) = p(w) > 1$, and $|f(\lambda w)| = |\lambda| p(w) = p(\lambda w)$, so we have a nice bound. Therefore, we can extend $f$ to an $F$ such that $F(w) > 1$ and $F(x) \le 1$ if $x \in C \subset C + B$. ⊠

Basically, the idea is the Minkowski functional; once you write that down, you're basically done.

Lecture 10.

## The Minkowski Functional and the Baire Category Theorem: 9/18/15

Last time, we had to rush through the Minkowski functional, so today we'll talk a little more about it. This is *not* a linear functional, but it does map into $\mathbb{F}$, so it's called a functional.

Specifically, given a nonempty $A \subseteq X$, where $X$ is an NLS, the Minkowski functional is defined as

$$p(x) = \inf\{t > 0 : x \in tA\},$$

which takes values in $[0, \infty]$. We then showed the following.

(1) If there's an open ball containing 0 and contained in $A$, then $p(x)$ is finite.
(2) $p$ is *positively homogeneous*, i.e. if $\lambda \ge 0$, then $p(\lambda x) = \lambda p(x)$.
(3) If $A$ is convex, then $p(x + y) \le p(x) + p(y)$. Well, we didn't actually show (3), so let's do that now. Suppose $x/r, y/s \in A$ (so that $r \ge p(x)$ and $s \ge p(y)$). By convexity,

$$\frac{x+y}{s+r} = \frac{r}{s+r}\frac{x}{r} + \frac{s}{s+r}\frac{y}{s} \in A,$$

and therefore $s + r \ge p(x + y)$. Since this is true for all such $s$ and $r$, passing to their infimum replaces them with $p(x)$ and $p(y)$, so $p(x) + p(y) \ge p(x + y)$.
(4) We did show that if $A$ is balanced, then $p$ is a seminorm.

This was sufficient to prove Lemma 9.6, but we have one more separating theorem to prove. This time, we don't need sets to be balanced, but we will require convexity.

**Lemma 10.1** (Separating hyperplane theorem). *Let $A$ and $B$ be disjoint, nonempty, convex subsets of an NLS $X$.*

*(1) If $A$ is open, then there exists an $f \in X^*$ and a $\gamma \in \mathbb{R}$ such that* $\mathrm{Re}(f(x)) \le \gamma \le \mathrm{Re}(f(y))$ *for all $x \in A$ and $y \in B$.*
*(2) If $A$ and $B$ are open, the above inequality is strict.*
*(3) If $A$ is compact and $B$ is closed, then the above inequality is also strict.*

*Proof.* We'll prove part 1; the others are similar. Moreover, it suffices to prove it for real fields, because if $\mathbb{F} = \mathbb{C}$, then we can view $X$ as a real vector space and get a real linear functional $g$ that satisfies the lemma over $\mathbb{R}$. Then, $f(x) = g(x) - ig(ix)$ satisfies the lemma for $\mathbb{C}$.

All right, so $\mathbb{F} = \mathbb{R}$, and $A$ is open and both are convex. We'll have to put the Minkowski functional into this proof somehow, so let's start by picking an $x \in A$ and a $y \in B$. Let $A - x = \{t - x : t \in A\}$, and define $B - y$ similarly. Then, let $C = (A - x) - (B - y) = \{t - s - x + y : t \in A, s \in B\}$, and for convenience, let $w = y - x$. We'll want to construct a Minkowski functional on $C$.

$C$ is open, since $A$ is; convex, because $A$ and $B$ are; and contains 0 (because we've moved $x$ and $y$ to the origin). But $w \notin C$, since $A$ and $B$ are disjoint. Let $Y = \mathbb{R}w$ and $g(tw) = t$; then, our Minkowski functional is $p(x) = \inf\{t > 0 : x \in tC\}$, which is well-defined and sublinear, and satisfies $p(w) \ge 1 = g(w)$, so $g(w) \le p(w)$, and therefore for any $y \in Y$, $g(y) \le p(y)$: if $\lambda \ge 0$, this follows from the positive homogeneity of $p$ and the linearity of $g$, and if $\lambda < 0$, $-\lambda g(w) \ge -\lambda p(w)$, and therefore $-g(-\lambda w) \le p(-\lambda w)$.

Thus, we can extend $g$ to $X$, and $g(x) \le 1$ on $C$, since $g(x) \le p(x)$ everywhere on $X$, and therefore $g(x) \ge -1$ for $x \in -C$, and so $|g(x)| \le 1$ on $C \cap (-C)$. Since this contains a neighborhood of the origin, $g$ is bounded, so $g \in X^*$.

If $a \in A$ and $b \in B$, then $a - b + w \in C$, and therefore $1 \ge g(a - b + w) = g(a) - g(b) + 1$, so $g(b) \ge g(a)$. Let $\gamma = \sup g(a)$ (or $\gamma = \inf g(b)$), and we're done. ⊠

This concludes our discussion of the Hahn-Banach theorem and its applications.

**The Open Mapping Theorem.** The open mapping theorem, which is the next major result for Banach spaces, helps us characterize what linear functionals can look like.

The following theorem is important in its own right, but we'll use it as an ingredient in the proof of the open mapping theorem.

**Theorem 10.2** (Baire category theorem). *Let $(X, d)$ be a complete metric space, and let $\{V_j\}_{j=1}^{\infty}$ be a sequence of open, dense subsets of $X$. Then, $V = \bigcap_{j=1}^{\infty} V_j$ is dense.*

In other words, open dense sets aren't exactly thin: they're actually surprisingly fat, so fat that a countable intersection of them is still fat, in a sense.

*Proof.* Let $W \subseteq X$ be any nonempty open set. Then, we have to show that $V \cap W \neq \emptyset$, since $V$ being dense is equivalent to intersecting every nonempty open.

Since $V_1$ is dense, then $W \cap V_1 \neq \emptyset$, so there's an $x_1 \in W \cap V_1$, and since $W$ is open, there's a $\overline{B_{r_1}(x_1)} \subseteq W \cap V_1$ (we have an open neighborhood, and can take the closure of a smaller ball); also, we can without loss of generality take $0 < r_1 < 1$, by shrinking $r_1$ if necessary.

In the same way, $V_2$ is open and dense and $B_{r_1}(x_1)$ is a nonempty open set, so there exists an $x_2 \in V_2 \cap B_{r_1}(x_1)$ and an $r_2 \in (0, 1/2)$, such that $\overline{B_{r_2}(x_2)} \subseteq B_{r_1}(x_1) \cap V_2$. Then, we can continue in this way, choosing for each $n$ an $x_n$ and an $r_n$ such that $\overline{B_{r_n}(x_n)} \subseteq B_{r_{n-1}}(x_{n-1}) \cap V_n$ and $0 < r_n < 1/n$.

Now, consider the sequence $\{x_n\}_{n=1}^{\infty}$, which is Cauchy, because if $i, j \ge n$, then $x_i, x_j \in B_{r_n}(x_n)$ and therefore $d(x_i, x_j) < 2/n$. Since $X$ is complete, this sequence converges to some $x \in X$. Since $x_i \in B_{r_n}(x_n)$ for all $i > n$, then $x \in \overline{B_{r_n}(x_n)}$, so $x \in V_n$ for all $n$. Since $x \in \overline{B_{r_1}(x_1)} \subseteq W$, then $x \in W \cap V$, so the intersection is nonempty, and thus $V$ is dense. ⊠

Some of you may have been disappointed to see that no category theory appeared in the statement or proof; in this part of mathematics, "category" has a different definition.

**Definition.** Let $(X, d)$ be a metric space.

- $A$ is *nowhere dense* if it has empty interior: $(\overline{A})^0 = \emptyset$.
- $A$ is *first category* if it can be written as a countable union of nowhere dense sets.
- If $A$ isn't first category, then it is called *second category*.

Using these definitions, the Baire category theorem says that a complete metric space is second category.

**Corollary 10.3.** *A complete metric space is not the countable union of nowhere dense sets.*

In other words, a complete metric space is fatter than that.

*Proof.* Suppose $X$ is such a union: $X = \bigcup_{j=1}^\infty M_j$ with each $M_j$ nowhere dense. Without loss of generality, each $M_j$ is closed (or just take their closures, which still cover $X$). Thus, by de Morgan's law, $\emptyset = \bigcap_{j=1}^\infty (M_j)^c$. Since $M_j$ is closed and nowhere dense, then $M_j^c$ is a dense open set, and therefore $\emptyset$ is the countable intersection of dense open sets, which contradicts Theorem 10.2. ⊠

Next time, we'll return to the world of Banach spaces.

Lecture 11.

## The Open Mapping Theorem: 9/21/15

Last time, we learned about the Baire category theorem. Today, we'll use it to prove the open mapping theorem.

**Definition.** A continuoue map $f : X \to Y$ is *open* if it maps open sets to open sets, i.e. if $U \subseteq X$ is open, then $f(U) \subseteq Y$ is open.

An arbitrary continuous map is not open; for example, $T : \mathbb{R}^2 \to \mathbb{R}^2$ sending $(x, y) \mapsto (x, 0)$ is perfectly continuous, but the image of $B_1(0)$ is $(0, 1) \times \{0\}$, which isn't open in $\mathbb{R}^2$.

In the infinite-dimensional case, things can become more interesting; for example, $T : \ell^2 \to \ell^2$ sending $e_n \to (1/n)e_n$ isn't open (the image of the unit ball isn't open), but is linear and surjective; the discrepancy is that this $T$ isn't bounded.

**Theorem 11.1** (Open mapping)**.** *Let $X$ and $Y$ be Banach and $T : X \to Y$ be a bounded, linear surjection. Then, $T$ is an open map.*

A bounded linear map is typical in this class; the key hypothesis in this theorem is that $T$ is surjective; the example $(x, y) \mapsto (x, 0)$ shows that this is important.

This is a pretty fundamental theorem about Banach spaces.

*Proof.* It suffices to show that $T(B_1(0))$ contains a $B_r(0)$ for some $r > 0$: if $U \subset X$ is open, then to check that $T(U)$ is open, we can pick a $y \in T(U)$ and a preimage $x$ (i.e. $T(x) = y$). Then, we can look at $U - x$, and since $T$ is linear, then $T(U - x) = T(U) - y$. But since $x$ and $y$ are now sent to the origin, we just need to pick a neighborhood of $x$ and make sure its image contains a neighborhood of $y$.

This is the proper way to think about the theorem: if you know what a bounded linear map looks like at the origin, you know what it looks like everywhere.

Since $T$ is onto, then we can write

$$Y = \bigcup_{j=1}^\infty T(B_j(0)).$$

Since $Y$ is a complete metric space, then the Baire category theorem tells us it's not the union of nowhere dense sets. Thus, there's some $k$ such that $T(B_k(0))$ isn't nowhere dense, i.e. there's an open $W_1 \subset \overline{T(B_k(0))}$. Thus, we can scale: $(1/2k)W \subseteq (1/2k)\overline{T(B_k(0))} = \overline{T(B_{1/2}(0))}$.

Since $W_1$ is open, there's a $y_0 \in Y$ and an $r > 0$ such that $B_r(y_0) \subseteq W \subseteq \overline{T(B_{1/2}(0))}$. This is almost everything:

$$\begin{aligned} B_r(0) &= B_r(y_0) - y_0 \\ &\subseteq B_r(y_0) - B_r(y_0) \\ &\subseteq \overline{T(B_{1/2}(0))} + \overline{T(B_{1/2}(0))} \\ &\subseteq \overline{T(B_1(0))}, \end{aligned}$$

by the triangle inequality. We'd be done, except that we had to take the closure (which ultimately came from the Baire category theorem). Thus, we'll show that if $\varepsilon > 0$, then $T(B_{1+\varepsilon}(0)) \supseteq B_r(0)$, because then

$$T(B_1(0)) = \frac{1}{1+\varepsilon} T(B_{1+\varepsilon}(0)) \supseteq B_{r/(1+\varepsilon)}(0).$$

Then, we won't need the closure anymore. Note that this isn't obvious, even if it seems obvious in the finite-dimensional case.

Fix a $y \in B_r(0)$ and an $\varepsilon > 0$. We know that $T(B_1(0)) \cap B_r(0)$ is dense in $B_r(0)$ (since we showed already its closure contains $B_r(0)$), so we can pick an $x_1 \in B_1(0)$ such that $\|y - Tx_1\| \le \varepsilon/2$.

Inductively, when $n \ge 1$, suppose we've picked $x_1, \dots, x_n$ such that $\|x_1\| \le 1$ and $\|x_j\| \le 2^{-j+1}\varepsilon$ and $\|y - T(x_1 + \cdots + x_n)\| < 2^{-n}\varepsilon r$. Let $z = y - T(x_1 + \cdots + x_n)$, so that $z \in B_{2^{-n}\varepsilon r}(0)$. Since $T(B_1(0)) \cap B_r(0)$ is dense in $B_r(0)$, we can scale things: there's an $x_{n+1} \in B_{2^{-n}\varepsilon}(0)$ such that $\|z - Tx_{n+1}\| \le 2^{-(n+1)}\varepsilon r$; thus, $\|y - T(x_1 + \cdots + x_{n+1})\| \le 2^{-(n+1)}\varepsilon r$.

Since the terms get smaller and smaller, $\sum_{j=1}^n x_j$ is a Cauchy sequence, so since $X$ is complete, then this sum converges to a point $x \in X$, such that

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

and $T$ is continuous, then $Ts_n \to Tx = y$. ⊠

The first part, showing it's true for $\overline{T(B_1(0))}$, is pretty easy, but then getting just one more $\varepsilon$ is surprisingly fussy.

**Corollary 11.2.** *If $X$ and $Y$ are Banach spaces and $T : X \to Y$ is a bounded, linear bijection (one-to-one and onto), then the inverse map exists and is a bounded linear functional.*

In other words, the inverse of a bounded linear functional is bounded linear. This is nice, and very useful.

*Proof.* It's easy to show $T^{-1}$ is linear. $T$ is open, so it takes open sets to open sets, and therefore for $T^{-1}$, the preimage of every open set is open, so $T^{-1}$ is continuous. ⊠

We now know enough to make the following definition.

**Definition.** If $X$ and $Y$ are Banach spaces, we say that they're *isomorphic as Banach spaces* if there exists a linear, bounded bijection $T : X \to Y$. If in addition $T$ preserves the norm, it's called a *isometry*.

This means that $X$ and $Y$ have the same vector space structure (since there's a bijective linear map) and same topology (there's a homeomorphism). If $T$ isn't an isometry, then the norms may be different, but they'll be equivalent, so $X$ and $Y$ are basically the same.

There's a closely related result about graphs of maps; we could have proven this first and used it to derive the open mapping theorem, though we'll go about it in the other direction.

**Definition.** Let $X$ and $Y$ be topological spaces, $D \subseteq X$ and $f : D \to Y$. Then, the *graph* of $f$ is $\text{graph}(f) = \{(x, f(x)) : x \in D\} \subseteq X \times Y$.

$X \times Y$ is where we're used to drawing graphs (such as $X, Y = \mathbb{R}$); we chose $D$ because the function might not be defined everywhere.

**Proposition 11.3.** *Let $X$ be a topological space, $Y$ be a Hausdorff space, and $f : X \to Y$ be continuous. Then,* $\text{graph}(f)$ *is closed in $X \times Y$.*

In the case of graphs we're most familiar with, this makes sense, as it's how we're used to thinking of continuity intuitively.

*Proof.* Let $U = X \times Y \setminus \text{graph}(f)$, and let $(x_0, y_0) \in U$, so $f(x_0) \le y_0$.

Recall that since $Y$ is Hausdorff, we can choose open neighborhoods $V, W \subset Y$ around $y_0$ and $f(x_0)$, respectively, that don't intersect. Then, we're going to consider $f^{-1}(W) \subset X$; specifically, $f^{-1}(W) \times V$ doesn't intersect $\text{graph}(f)$, and is an open neighborhood of $(x_0, y_0)$. ⊠

If that proof didn't make sense, drawing a picture will likely help.

**Definition.** Let $X$ and $Y$ be NLSes, $D \subseteq X$ be a linear subspace, and $T : D \to Y$ be linear. Then, we say that $T$ is a *closed* operator if $\text{graph}(T)$ is closed in $X \times Y$.

An open map takes open sets to open sets; a bounded map takes bounded sets to bounded sets; but a closed operator *doesn't* take closed sets to closed sets. This can be confusing.

If $X$ and $Y$ are metric spaces and $f : D \to Y$ is continuous, then $\text{graph}(f)$ being closed means that if $\{x_n\}_{n=1}^{\infty} \subseteq D$ such that $x_n \to x$ and $Tx_n \to y$, then $x \in D$ and $y = Tx$.

**Theorem 11.4** (Closed graph theorem)**.** *Let $X$ and $Y$ be Banach spaces, and $T : X \to Y$ be linear. Then, $T$ is bounded iff it's closed.*

*Proof.* The forward direction is true in general (continuous implies closed).

In the other direction, suppose graph($T$) is closed, and therefore is a closed linear subspace of $X \times Y$. This is a very important point: since $X$ and $Y$ are Banach, then $X \times Y$ is Banach, and since graph($T$) is closed in it, then graph($T$) is also a Banach space, with the *graph norm*[9]

$$\|(x, Tx)\| = \|x\|_X + \|Tx\|_Y.$$

Define two projection operators $\pi_1 : (x, y) \mapsto x$ and $\pi_2 : (x, y) \mapsto y$, so that we have maps

$$\begin{tikzcd} & \operatorname{graph}(T) \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ X & & Y. \end{tikzcd}$$

$\pi_1$ is a linear bijection between Banach spaces, and is bounded (by the triangle inequality, $\|x\|_X \le \|(x, Tx)\| + \|Tx\|_Y$), so by Corollary 11.2, its inverse $\pi_1^{-1}$ is a bounded linear functional. Moreover, $\pi_2$ is bounded linear for the same reasons, so $T = \pi_2 \circ \pi_1^{-1}$ is bounded as well. ⊠

Lecture 12.

## The Uniform Boundedness Principle: 9/23/15

Recall that last time, we proved the closed graph theorem, Theorem 11.4, which states that if $X$ and $Y$ are Banach spaces and $T : X \to Y$ is linear, then $T$ has a closed graph iff it's bounded.

**Corollary 12.1.** *Let $X$ and $Y$ be Banach spaces, $D \subset X$ be a subspace, and $T : D \to Y$ be closed. Then, $T$ is bounded iff $D$ is closed.*

*Proof.* In the reverse direction, if $D$ is closed, then it's Banach, so we can apply Theorem 11.4.

Conversely, suppose $T$ is bounded, and let $\{x_n\}_{n=1}^\infty \subseteq D$ such that $x_n \to x$ in $X$. Since $T$ is bounded and linear, then $\{Tx_n\}_{n=1}^\infty$ is Cauchy.[10]

Since $Y$ is complete, then $Tx_n \to y$ for some $y \in Y$. And since $T$ is a closed operator, then graph($T$) = $\{(x, Tx)\}$ is closed. Since $x_n \to x$ and $Tx_n \to y$, then $(x, y) \in$ graph($T$), so $y = Tx$ and thus $x \in D$. Therefore $D$ contains its limit points, and so is closed. ⊠

Hopefully this illustrates some uses of the closed graph theorem.

**Example 12.2.** Though continuous implies closed, the converse isn't true. Here's an example. Let $X = C([0, 1])$, the continuous functions with the $L^\infty$ norm, and let $D = C^1([0, 1]) \subset X$, the $C^1$ functions. Let $T : D \to X$ be the derivative operator, so $T(f) = f'$.

This is perfectly well defined: if $f$ is $C^1$, then $f'$ is continuous. Then, $D \neq X$ (e.g. $f(x) = |x - 1/2|$), but $\overline{D} = X$, so $D$ is *not* closed in $X$ (intuitively, any continuous but not differentiable function can be well approximated by a $C^1$ function).

- First, we'll see that $T$ isn't continuous (equiv. bounded). $T(x^n) = nx^{n-1}$, but $\|x_n\| = 1$ and $\|Tx_n\| = n \to \infty$.
- However, it is closed. Let $\{f_n\}_{n=1}^\infty \subseteq D$ have a limit $f_n \to f$ in $X$, and such that $f_n' \to g$ in $X$. We want to show that $g = f'$. This follows from the fundamental theorem of calculus: for each $n$,

$$f_n(t) = f_0(t) + \int_0^t f_n'(\tau)\, d\tau.$$

[9]Though this was defined in a way mirroring the $\ell^1$ norm, you can use the analogous definition with any $\ell^p$ including $\ell^\infty$, since, as we proved on a problem set, they're all equivalent.

[10]This is because $\|Tx_n - Tx_m\| = \|T(x_n - x_m)\| = \|T\|\|x_n - x_m\| \to 0$.

Since convergence in $L^\infty$ implies pointwise convergence. Then, by the dominated convergence theorem, these integrals also converge (recall that continuous functions on compact sets are bounded), so

$$f(t) = f(0) + \int_0^t g(\tau)\,\mathrm{d}\tau = f(0) + \int_0^\tau f'(\tau)\,\mathrm{d}\tau.$$

Thus, $f'(t)$ exists and $g = f'$, so $f \in C^1([0,1])$.

Continuous is better than closed, but closed is part of the way there, in some sense.

We've talked about two of the three important theorems about NLSes: the Hahn-Banach theorem and the open mapping theorem. Here's the third.

**Theorem 12.3** (Uniform boundedness principle). *Let $X$ be Banach, $Y$ be an NLS, and $\{T_\alpha\}_{\alpha\in I} \subseteq B(X,Y)$. Then, one of the following holds.*

*(1) The collection is uniformly bounded: there's an $M$ such that $\sup_{\alpha\in I}\|T_\alpha\|_{B(X,Y)} \le M$,*
*(2) There's a point where they're not: there's an $x \in X$ such that $\sup_{\alpha\in I}\|T_\alpha X\| = \infty$.*

The point is, if these functions aren't uniformly bounded, then they all blow up at a given point, when *a priori* they could do so in different places. This is true no matter how large the collection $I$ is; it could very well be uncountable.

The proof in the notes is nice, but a little fussy to prove, and uses the Baire category theorem. We'll give a proof based on the following lemma, which is a little nicer.

**Lemma 12.4.** *Let $X$ and $Y$ be NLSes and $T : X \to Y$ be a bounded linear map. For any $x \in X$ and $r > 0$, $\sup_{y\in B_r(x)}\|Ty\| \ge r\|T\|$.*

The idea is if $x = 0$, we have equality, but even if we don't, this is still a one-sided bound.

*Proof.* For a visualization, it may help to think about the case $X = Y = \mathbb{R}$, where graph$(T)$ is a line with slope $\|T\|$ through the origin. Here, $\sup_{y\in B_r(x)}\|Ty\| = \|T\|(\|x\|+r) \ge \|T\|r$.

More generally, we'll think of the "larger" and "smaller" parts of $B_r(x)$. The triangle inequality tells us that since $z = (1/2)(x+z) - (x-z)$, then

$$\begin{aligned}\|Tz\| &\le \frac{1}{2}(\|T(x+z)\| + \|T(x-z)\|)\\ &\le \max\{\|T(x+z)\|, \|T(x-z)\|\}.\end{aligned}$$

If we take the supremum over $z \in B_r(0)$, we know that

$$r\|T\| \le \sup_{z\in B_r(0)}\|T(x+z)\| = \sup_{y\in B_r(x)}\|Ty\|.$$ ⊠

This is a nice geometric result, and relatively easy to prove. Then, we'll use it to attack the uniform boundedness principle.

*Proof of Theorem 12.3.* Since we want to show one of two things is true, let's assume $\sup_\alpha\|T_\alpha\| = \infty$. Choose a countable subcollection $\{T_n\}_{n=1}^\infty$ such that $\|T_n\| \ge 4^n$. Then, choose $x_0 = 0 \in X$ and choose $x_n \in X$ such that $\|x_n - x_{n-1}\| \le 3^{-n}$ (so that $x_n \in B_{3^{-n}}(x_{n-1})$), so that by Lemma 12.4, $\|T_n x_n\| \ge (2/3)3^{-n}\|T_n\|$.

Since $\{x_n\}$ is Cauchy, then it converges to some $x \in X$, and $\|x - x_n\| \le (1/2)3^{-n}$, since

$$\begin{aligned}\|x - x_n\| &= \lim_{m\to\infty}\|x_m - x_n\|\\ &\le \lim_{m\to\infty}\sum_{j=m}^{n+1}\|x_j - x_{j-1}\|\\ &\le \lim_{m\to\infty}\left(3^{-m} + 3^{-(m-1)} + \cdots + 3^{-(n+1)}\right)\\ &\le \lim_{m\to\infty} 3^{-n}\sum_{k=1}^\infty\left(\frac{1}{3}\right)^k = \frac{1}{2}3^{-n},\end{aligned}$$

since it's a nice old geometric series.

Finally, we're going to look at $T_n x$.

$$\|T_n x\| \le \|T_n(x - x_n)\| + \|T_n x_n\| \le \|T_n\|\|x - x_n\| + \|T_n x\|.$$

Thus, we know that

$$\frac{2}{3}3^{-n}\|T_n\| \leq \|T_n\|\frac{1}{2}3^{-n} + \|T_nx\|,$$

and therefore

$$\|T_nx\| \geq \frac{1}{6}3^{-n}\|T_n\| \geq \frac{1}{3}\left(\frac{4}{3}\right)^n,$$

which goes to infinity. ⊠

**The Double-Dual.** If $X$ is an NLS, then $X^* = B(X,\mathbb{F})$ is a Banach space, and we can form the *double-dual* $X^{**} = B(X^*,\mathbb{F})$, which is also a Banach space. It's possible to interpret $X$ as sitting inside $X^{**}$.

For any $x \in X$, define the evaluation map $E_x : X^* \to \mathbb{F}$ by $E_x(f) = f(x)$: that is, we evaluate $f$ at $x$. Since $f : X \to \mathbb{F}$, then $E_x$ is well-defined, and it's linear: if $f, g \in X^*$ and $\lambda \in \mathbb{F}$, then

$$E_x(f+g) = (f+g)(x) = f(x) + g(x) = E_x(f) + E_x(g)$$
$$E_x(\lambda f) = (\lambda f)(x) = \lambda(f(x)) = \lambda E_x(f).$$

$E_x$ is also bounded, which is more interesting. If you think about what the norm means, then

$$\|E_x\|_{X^{**}} = \sup_{\substack{f\in X^*\\ f\neq 0}} \frac{|E_x(f)|}{\|f_{X^*}\|} = \sup_{\substack{f\in X^*\\ f\neq 0}} \frac{f(x)}{\|f\|_{X^*}} = \|x\|_X,$$

where the last equality is due to Corollary 8.5.

**Definition.** Let $(M,d)$ and $(N,\rho)$ be metric spaces. Then, $f : M \to N$ is an *isometry* if $\rho(f(x), f(y)) = d(x,y)$ for all $x, y \in M$. If $f$ is surjective, $M$ and $N$ are said to be *isometric*.

Note that $f$ is always injective, because if $f(x) = f(y)$, then $\rho(f(x), f(y)) = 0 = d(x,y)$, so $x = y$. Thus, isometric spaces are given by a bijection $f$.

Anyways, that's what's going on here: not only is the metric the same, but the norm is the same. We have a map $E : X \to X^{**}$ sending $x \mapsto E_x$. $E$ is a bounded linear map, and an isometry. Therefore, $\widetilde{X} = \{E_x \in X^{**} : x \in X\} \subset X^{**}$ is isomorphic and isometric to $X$. It might not be all of $X^{**}$, but we've embedded $X$ into its double-dual.

**Definition.** Sometimes, $X = X^{**}$ (i.e. $\widetilde{X} = X^{**}$). If this is true, $X$ is said to be *reflexive*.

In general, $X \subseteq X^{**} \subseteq X^{****} \subseteq \cdots$; if $X = X^{**}$, then this entire chain collapses to equalities. Similarly, we could have started with $X^*$ and $X^{***}$, and so on.

**Theorem 12.5.** *If $X$ is reflexive, then $X^*$ is reflexive.*

This is left to the exercises, but isn't hard to prove. Notice, however, that the converse isn't true.

**Example 12.6.** If $1 \leq p < \infty$, we know that $(\ell^p)^* = \ell^q$, so $\ell^p$ is reflexive for $1 < p < \infty$. $\ell^1$ and $\ell^\infty$ aren't as nice, so we don't have reflexivity. (The duality follows from something on the homework this week.)

Since $(L^p(\Omega))^* = L^q(\Omega)$ for $1 < p < \infty$, then $L^p(\Omega)$ is reflexive (which follows from the Radon-Nikodym theorem). Similarly, $L^1(\Omega)$ and $L^\infty(\Omega)$ are more complicated.