14 Banach spaces

> "Banach didn't just develop Banach spaces for the sake of it. He wanted to put many spaces under one heading. Without knowing the examples, the whole thing is pointless. It is a mistake to focus on the techniques without constantly, not only at the beginning, knowing what the good examples are, where they came from, why they're done. The abstractions and the examples have to go hand in hand.
>
> You can't just do particularity. You can't just pile one theorem on top of another and call it mathematics. You have to have some architecture, some structure. Poincaré said science is not just a collection of facts any more than a house is a collection of bricks. You have to put it together the right way. You can't just list 101 theorems and say this is mathematics. That is not mathematics."

(Sir Michael Atiyah, from `http://www.johndcook.com/blog/2013/09/24/interview-wi`

Definition 14.1. A Banach space is a complete normed space.

Examples
- $\mathbb{C}^n$ with the l^p norm.

We will mostly, but not exclusively, be interested in infinite dimensional spaces.
- Any Hilbert space.
- $C(X)$ for any compact metric or topological space X, with the sup norm.
- $L^p(X, \mu)$ for any measure space (X, μ).
- $M(X)$ for any measurable space $(X, \mathcal{M})$.
- The space of BLTs on a Hilbert space H.
- Sobolev spaces.
- Hölder spaces.

14.1 Basic Concepts

• Let X be a Banach space with norm $\|\cdot\|_X$. We say that the norm $\|\cdot\|'$ on X is equivalent to $\|\cdot\|_X$ if there exists $C > 0$ such that
$$C^{-1}\,\|\cdot\|_X \le \|\cdot\|' \le C\,\|\cdot\|_X\,.$$

Theorem 14.2. *Let V be a finite dimensional vector space. Then any two norms are equivalent.*

Proof: Choose any basis $v_1, \ldots, v_n$ of V and define a norm by
$$\|\sum_{i=1}^n a_i v_i\|_1 = \sum_{i=1}^n |a_i|.$$
It is straightforward to check that this is a complete norm.

Now given any other norm $\|v\|$, we have
$$\|\sum_{i=1}^n a_i v_i\| \le \sum_{i=1}^n |a_i|\|v_i\| \le M\|\sum_{i=1}^n a_i v_i\|_1,$$
where $M = \max \|v_i\|$. Thus the function $v \mapsto \|v\|$ is continuous with respect to the norm $\|\cdot\|_1$. Therefore it is bounded. It's also bounded away from zero on the unit 'sphere' S of vectors with $\|v\|_1 = 1$, simply because it is nonzero at each point on S and S is compact with respect to $\|\cdot\|_1$. Thus there exists C such that
$$v \in S \implies C^{-1} \le \|v\| \le C.$$
Homogeneity then implies that $\|\cdot\|$ is equivalent to $\|\cdot\|_1$. But then, any two norms are equivalent, since they are both equivalent to $\|\cdot\|_1$. □

There is a converse: if X is infinite dimensional, then there exists inequivalent norms on X.

• We say that X is separable if X has a countable dense subset. (Note that while amongst Hilbert spaces, the separable ones are the most common and the most useful, when we study Banach spaces we often encounter interesting non-separable examples, such as ℓ^∞.)

• We define linear transformations, and bounded linear transformations (BLTs) just as in the Hilbert space case. We denote by $B(X, Y)$ the set of BLTs from the Banach space X to the Banach space Y. As before, for linear transformations, boundedness and continuity are equivalent. Also as in the Hilbert space setting, we define the operator norm for $T \in B(X, Y)$
$$\|T\| = \sup_{\|x\|_X = 1} \|Tx\|_Y$$
and this endows $B(X, Y)$ itself with a Banach space structure. In the case that $Y = \mathbb{C}$, we call $B(X, \mathbb{C})$ the dual space of X, and denote it X^*. Elements of X^* are called continuous linear functionals. Note that, a priori, for an abstract Banach space X it is not obvious that X^* contains any nonzero elements (that it does so is the content of the Hahn-Banach theorem).

If $T : X \to Y$ is a BLT with a two-sided inverse which is a BLT from Y to X, then we say that T is an isomorphism between X and Y, and that X is isomorphic to Y. We say that T is an

isometric isomorphism if it is an isomorphism which is norm preserving, i.e. if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$.

• Given two Banach spaces X and Y, we define the product Banach space $X \times Y$ to be the vector space $X \times Y$ with norm $\|(x, y)\| = \|x\|_X + \|y\|_Y$.

• Let $Z \subset X$ be a closed subspace. Then it is a Banach space in its own right with the same norm (restricted to Z). This is simply because any closed subset of a complete metric space is also complete. We can also define a quotient Banach space, on the vector space X/Z, via the norm
$$\|Z + x\| = \inf_{z \in Z} \|x + z\|_X.$$
It is true, but not obvious, that this norm is complete on X/Z. Note that the quotient map $X \to X/Z$ is bounded with norm 1.

• There is in a general Banach space no notion of 'orthogonal complement' of a closed subspace Z, in the sense of a canonical choice of closed subspace W such that X is isomorphic to $Z \oplus W$.

• We say that a closed subspace $Z \subset X$ is *complemented* if there exists a closed subspace W such that every $x \in X$ has a unique representation $x = z + w$, with $z \in Z$ and $w \in W$. This condition on Z and W turns out to be equivalent to the apparently stronger one of asking that X be isomorphic to $Z \oplus W$ via the natural map $(z, w) \to z + w$. the natural map $W \to X/Z$ being an isomorphism (see example below).

• In general, closed subspaces of Banach spaces are not complemented (unlike for Hilbert spaces). A concrete example is $X = L^1(S^1)$, Z the subspace of functions with zero Fourier coefficients f_n for $n < 0$ (see Rudin, Functional Analysis).

14.2 Baire category and its consequences

There is a group of major theorems that use the Baire 'category' theorem. First recall that a subset of a metric space is *nowhere dense* if its closure has empty interior.

(E.g. $\mathbb{R} \subset \mathbb{R}^2$ is nowhere dense. The complement of a dense open set is nowhere dense.)

Theorem 14.3 (Baire)**.** *Let X be a complete metric space. Then the complement of a countable union of nowhere dense sets is dense.*

Proof: Let A_n be sequence of nowhere dense sets, U an arbitrary open set in X. It suffices to show that U intersects $(\cup_i A_i)^c$.

Since $\overline{A_1}$ has no interior, $(\overline{A_1})^c$ intersects U. So we can choose $x_1 \in U \setminus \overline{A_1}$ and an $\epsilon_1 > 0$ such that $\overline{B_1} = \overline{B(x_1, \epsilon_1)} \subset (U \setminus \overline{A_1})$. Let $U_2 = B(x_1, \epsilon_1/2)$. Now inductively, given an open set U_i, choose[1] an open ball B_i with closure in U_i but disjoint from $\overline{A_i}$, and let U_{i+1} be the concentric open ball of half the radius. Then we have constructed a sequence of balls (B_n) such that
$$\overline{B_n} \cap \bigcup_1^n A_i = \emptyset$$
with radii converging to 0. Then $\bigcap_i B_i$ is a singleton which is in none of the A_n. Thus $U \setminus \bigcup A_n \neq \emptyset$. □

The Baire category theorem has a number of useful corollaries.

Corollary 14.4. *The intersection of a sequence of dense open sets in a complete metric space is dense.*

Corollary 14.5. *If a complete metric space is the union of an increasing family of closed sets, then one of them must have nonempty interior.*

Corollary 14.6. *In particular, an infinite dimensional Banach space cannot have countable (algebraic) dimension. For given (x_n) linearly dependent, $X_n = \mathrm{span}\{x_1, \ldots, x_n\}$ is closed for each n and has empty interior. Thus $\bigcup X_n \subsetneq X$ is a proper subset.*

Definition 14.7. A set S in a topological space X is *meagre*, or *of the first category* if S lies in the union of countable many nowhere dense sets. Otherwise S is *non-meagre* or *of the second category.*

So meagre sets are 'small' in a certain sense, non-meagre ones 'large'. Beware: one can decompose $\mathbb{R} = A \cup B$ where A is meagre and of full measure, B is non-meagre and of measure zero!

Theorem 14.8 (Principle of uniform boundedness)**.** *Let X, Y be Banach spaces, and let $\{T_\alpha\} \subset B(X, Y)$ be a family of BLTs which is pointwise bounded, that is, for each $x \in X$,*
$$\sup_\alpha \|T_\alpha x\| < \infty.$$
Then $\{\|T_\alpha\|\}$ is bounded, that is,
$$\sup_\alpha \|T_\alpha\| < \infty\,.$$

[1] Observe the axiom of choice being used!

Proof: Define

$$S_k = \{x \in X : \|T_\alpha(x)\| \le k \text{ for all } \alpha\}\,.$$

The pointwise boundedness hypothesis is exactly that $\bigcup_k S_k = X$. By continuity of the T_α, each S_k is closed. Baire category implies that some S_k has interior. (Since $S_k = kS_1$ in this particular case, in fact they all do). So there is $x_0 \in S_k$ and $\delta > 0$ such that $B(x_0, \delta) \subset S_k$. Now for any $x \in B(0, \delta)$,

$$\|T_\alpha x\| \le \|T_\alpha x_0\| + \|T_\alpha(x - x_0)\| \le 2k.$$

This gives $\|T_\alpha\| \le 2k\delta^{-1}$ for all α. □

Corollary 14.9. *Let X, Y be Banach spaces, $(T_n) \subset B(X, Y)$ a sequence which converges pointwise, that is $Tx = \lim T_n x$ exists for each $x \in X$. Then $T \in B(X, Y)$.* □

Proof: It is straightforward to check that T is linear. Since the T_n converge pointwise, they are pointwise bounded. By the UBP, they are uniformly bounded in norm. But $\|Tx\| \le \sup \|T_n x\|$, hence $\|T\| \le \sup \|T_n\|$ so this implies that T is bounded. □

Example. Suppose that X, Y, Z are Banach spaces, and that $T : X \times Y \to Z$ is a bilinear map which is bounded as a function of each variable separately (with the other held fixed). Then T is bounded, i.e.

$$\|T(x, y)\|_Z \le M\|x\|_X\|y\|_Y.$$

(For example, if $X = Y = Z = L^1(\mathbb{R})$, then $(f, g) \mapsto f + g$ is a bounded linear transformation, while $(f, g) \mapsto f * g$ is a bounded bilinear transformation.)

Proof: the hypothesis is that there exist constants $C_1(x), C_2(y)$ such that

$$\|T(x, y)\| \le C_1(x)\|y\|, \quad \|T(x, y)\| \le C_2(y)\|x\|.$$

Now define maps $T_x(y) = T(x, y) : Y \to Z$ and $T_y(x) = T(x, y) : X \to Z$. Regard T_x as a family of maps parametrized by x such that $\|x\| \le 1$. Then $\|T(x, y)\| \le C_2(y)\|x\|$ implies that the family T_x is pointwise bounded for all $y \in Y$. We conclude that $\|T_x\| \le M$ for all $\|x\| \le 1$. By scaling we obtain $\|T(x, y)\| \le M\|x\|\|y\|$ as required. □

Example. There exists a continuous function whose Fourier series diverges at zero.

Proof: Consider the map Φ_n that takes $f \in C([0, 2\pi])$ to its nth partial Fourier sum, i.e.

$$\Phi_n(f) = \frac{1}{2\pi}\sum_{|j| \le n} e^{in\theta}\langle f, e^{in\theta}\rangle.$$

Φ_n is an integral operator with kernel ('Dirichlet kernel')

$$\begin{aligned}
K(\theta, \theta') &= \frac{1}{2\pi}\sum_{|j| \le n} e^{in(\theta - \theta')} \\
&= \frac{1}{2\pi}\frac{e^{i(n+1/2)(\theta-\theta')} - e^{-i(n+1/2)(\theta-\theta')}}{e^{i/2(\theta-\theta')} - e^{-i/2(\theta-\theta')}} \\
&= \frac{1}{2\pi}\frac{\sin((n + 1/2)(\theta - \theta')}{\sin(1/2)(\theta - \theta')}.
\end{aligned}$$

In particular, if we let ϕ_n be the linear functional $\phi_n(f) = (\Phi_n(f))(0)$, then

$$\phi_n(f) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\sin((n + 1/2)y}{\sin(1/2)y} f(y)\, dy.$$

Now, here are two facts which are not too hard to verify:

- For any linear functional
$$f \mapsto \int_0^{2\pi} K(y) f(y)\, dy$$
acting on $C([0, 2\pi])$, the norm is precisely
$$\int_0^{2\pi} |K(y)|\, dy.$$
- We have
$$\int_0^{2\pi} \left|\frac{\sin(n + 1/2)y}{\sin(1/2)y}\right| dy \ge C \log n$$
for large n.

It follows that the ϕ_n are not uniformly bounded. By UBP, there exists an $f \in C([0, 2\pi])$ on which the ϕ_n are not uniformly bounded. But that says that the Fourier series of f diverge at 0. □

You might appreciate how difficult it is to make this proof constructive!

14.3 Closed Graph Theorem

Suppose that X, Y are topological spaces, $f : X \to Y$ is continuous. Then $Gr(f) = \{(x, f(x)) : x \in X\}$ is necessarily a closed subset of the product $X \times Y$. Is the converse true? The answer is yes, provided that the graph is the graph of a function defined on the whole of X.

Example. Consider $Y = C[0, 1]$, X the subspace of continuously differentiable functions, both under the uniform norm, and $D : X \to Y : f \mapsto f'$. Suppose $(f_n, f_n') \to (g, h)$ in $X \times Y$.

Thus $f_n \to g$, $f_n' \to h$ both uniformly on $[0, 1]$. It is a classical result that this means $g' = h$. Thus $Gr(D)$ is closed. But since $D(t^n) = nt^{n-1}$, D is *not* continuous.

Definition 14.10. Let X, Y be normed spaces, $A \subset X$ a subspace and $T : A \to Y$ a linear mapping. Then T is *closed* if $(x_n) \subset A$, $x_n \to x$ and $Tx_n \to y$ implies $x \in A$ and $Tx = y$. That is, T is closed iff its graph is a closed subspace of $X \times Y$.

The example above shows that $D : (C^1[0, 1], \|\cdot\|_\infty) \to (C[0, 1], \|\cdot\|_\infty)$ is closed yet not continuous.

Theorem 14.11 (Closed Graph Theorem). *Let X, Y be Banach spaces, $T : X \to Y$ a closed linear transformation. Then T is continuous.*

The following result is an ingredient of the proof, as well as being useful for other applications.

Lemma 14.12. *Let X be Banach space, and $\lambda > 1$ fixed. Supose that $A \subset X$ has the property that for any $x \in X$, $\|x\| \le 1$, there exists $y \in A$ such that $\|x - y\| \le \lambda^{-1}$. Then for $\|x\| \le 1$ there is a sequence $(y_n) \subset A$ such that $x = \sum \lambda^{-n} y_n$.*

Proof: Take $\|x\| \le 1$, so there is $y_0 \in A$ with $\|x - y_0\| \le \lambda^{-1}$. Then $\|\lambda(x - y_0)\| \le 1$ so there is $y_1 \in A$ such that $\|\lambda(x - y_0) - y_1\| \le \lambda^{-1}$, that is, $\|(x - y_0) - \lambda^{-1}y_1\| \le \lambda^{-2}$. Inductively, there is $(y_n) \subset A$ with $\|x - \sum_{r=0}^n \lambda^{-r} y_r\| \le \lambda^{-n-1} \to 0$. □

Proof of the Closed Graph Theorem: For each k define

$$A_k = \{x \in X : \|Tx\| \le k\}\,.$$

Then $\bigcup A_k = X$, so some $\overline{A_j}$ has interior. Therefore there is an $x \in X$ and $\delta > 0$ such that $\overline{B(x, \delta)} \subset \overline{A_j}$. Then also $\overline{B(-x, \delta)} \subset \overline{A_j}$. In other words, any element of $\overline{B(x, \delta)}$ or of $\overline{B(-x, \delta)}$ can be approximated (to arbitrary accuracy) by elements of A_j. Since $x_1, x_2 \in A_j$ implies that $x_1 + x_2 \in A_{2j}$, this means that any element of $\overline{B(0, \delta)}$ can be approximated by elements of A_{2j}. Therefore, taking $M \ge 2j/\delta$, the closure of A_M contains $\overline{B(0, 1)}$.

We apply Lemma 14.12 with $A = A_M$, $\lambda = 2$. Thus for any x, $\|x\| \le 1$, we can write $x = \sum 2^{-n} y_n$ with $y_n \in A$ (and hence $\|Ty_n\| \le M$). Then

$$\|T\left(\sum_{r=m}^n 2^{-r} y_r\right)\| \le M\sum_{r=m}^n 2^{-r} \to 0\,,$$

so that $T\left(\sum_{r=0}^n 2^{-r} y_r\right)$ converges to some limit as $n \to \infty$. But T is closed, so

$$Tx = \lim_n T\left(\sum_{r=0}^n 2^{-r} y_r\right) = \sum_{r=0}^\infty 2^{-r} Ty_r\,,$$

so $\|Tx\| \le \sum M2^{-r} = 2M$. □

Corollary 14.13 (Bounded Inverse Theorem). *Let X, Y be Banach spaces, $T : X \to Y$ a continuous linear bijection. Then T^{-1} is continuous.*

Proof: T is closed , so T^{-1} is closed. □

Example. Suppose X contains two closed subspaces Y and Z, such that every $x \in X$ has a unique representation $x = y + z$ with $y \in Y$ and $z \in Z$. Then $\|x\|$ is comparable to $\|y\| + \|z\|$ as a norm on X. Indeed, $\|x\| \le \|y\| + \|z\|$ by the triangle inequality. To get the other inequality, consider the map from $Y \oplus Z$ to X given by $(y, z) \mapsto y + z$. This is a continuous bijection, so by the BIT the inverse is continuous. Hence $\|x\| \ge c(\|y\| + \|z\|)$.

An open map is an map A such that $A(U)$ is open whenever U is open. In a metric space, this is equivalent to the statement that for any open ball $B(x, \delta)$, there is some open ball $Tx \in B(Tx, \epsilon) \subset T(B(x, \delta))$.

Corollary 14.14 (Open Mapping Theorem, or 'Banach-Schauder' Theorem). *Let X, Y be Banach spaces, $T : X \to Y$ a continuous linear surjection. Then T is open.*

Proof: Since T is continuous, $\ker T$ is closed, and so $X/\ker T$ is a Banach space under the quotient norm. The induced map $\tilde{T} : X/\ker T \to Y$ is a continuous bijection, and therefore has a continuous inverse by the bounded inverse theorem (Theorem 14.13). Therefore it is an open map. But the quotient map $X \to X/\ker T$ is also open. Hence T is open. □

See also `http://www.math.sc.edu/~schep/Openmapping.pdf` for another discussion of the open mapping theorem.

Exercise. Show that the open mapping theorem fails if either X or Y is not complete.

14.4 Hahn-Banach theorem

The content of the Hahn-Banach theorem is essentially that there are a lot of continuous linear functionals on any Banach space. It is normally stated in a slightly more general form. A 'gauge function' on a real vector space X is a function $p : X \to \mathbb{R}$ such that

$$p(x+y) \le p(x) + p(y) \text{ and } p(\alpha x) = \alpha p(x)$$

for all $x, y \in X$ and $\alpha \ge 0$. Any norm on X is a gauge function. However, gauge functions need not be nonnegative. For example, any linear function is a gauge function; also, the sum of any gauge function and a linear function is another gauge function.

Theorem 14.15 (Hahn-Banach Theorem, version 1)**.** *Let p be a gauge function on a real vector space X, Y a subspace of X and $\phi : Y \to \mathbb{R}$ linear and satisfying $\phi(x) \le p(x)$ for $x \in Y$. Then there is a linear functional $\tilde{\phi} : X \to \mathbb{R}$ such that $\tilde{\phi}|_Y = \phi$ and $\phi(x) \le p(x)$ for $x \in X$.*

Proof: We first show that if Y' is the subspace spanned by Y and a single element $z_0 \in X \setminus Y$, then we can extend ϕ to Y' such that $\phi \le p$ on Y'. It is convenient to subtract the linear function $z + tz_0 \to \phi(z)$ (defined on Y') from both p and ϕ. Thus, we may assume wlog that $\phi = 0$ and $p \ge 0$ on Y.

Then since p is a gauge function, for any $z', z'' \in Y$, we have

$$p(z'' + z_0) + p(z' - z_0) \ge p(z'' + z') \ge 0.$$

Therefore, for arbitrary $z', z'' \in Y$ we have

$$p(z'' + z_0) \ge -p(z' - z_0). \tag{14.1}$$

Define $\beta = \inf_{z'' \in Y} p(z'' + z_0)$. This is finite since by (14.1), $-p(z' - z_0)$ is a lower bound for any $z' \in Y$, and moreover, $\beta \ge -p(z' - z_0)$ for any $z' \in Y$. Now define $\phi(z + tz_0) = t\beta$. Then for any $t > 0$, we have

$$\phi(z + tz_0) = t\beta \le tp(z/t + z_0) = p(z + tz_0).$$

On the other hand, for $t > 0$ we have

$$\phi(z - tz_0) = -t\beta \le tp(z/t - z_0) = p(z - tz_0).$$

Thus we have extended ϕ from Y to Y' with the subgauge property.

The result now follows by Zorn's Lemma. Let $\mathcal{P}$ be the set of pairs (Z, ψ) where Z is a subspace of X containing Y, and $\psi : Z \to \mathbb{R}$ extends ϕ and satisfies $\psi \le p$ on Z. Partially order $\mathcal{P}$ by

$$(Z_1, \psi_1) \le (Z_2, \psi_2) := Z_1 \subseteq Z_2 \text{ and } \psi_2|_{Z_1} = \psi_1\ .$$

Any chain (Z_α, ψ_α) in $\mathcal{P}$ has upper bound $(\bigcup Z_\alpha, \psi)$ where $\psi|Z_\alpha = \psi_\alpha$. Thus $\mathcal{P}$ has a maximal element (Z, ψ). Necessarily $Z = X$ or else we can extend as above. □

Next we need a version for complex Banach spaces.

Theorem 14.16 (Hahn-Banach theorem, version 2)**.** *Let X be a complex Banach space, $Y \subset X$ a subspace and $\phi : Y \to \mathbb{C}$ a bounded (complex)-linear functional on Y. Then ϕ extends to a bounded linear functional ϕ' on X with $\|\phi'\| = \|\phi\|$.*

Proof: The real part ϕ_r of ϕ is a real-linear functional on X, regarded as a real vector space. Let p be $\|\phi\|$ times the norm on X. Then p is a gauge function and $\phi_r \le p$ on Y. Extend ϕ_r to X using Hahn-Banach version 1. Then define $\phi_i(x) = \phi_r(-ix)$ and define $\phi' = \phi_r + i\phi_i$. Then ϕ' is complex linear, agrees with ϕ on Y and has the same norm as ϕ, since if $\arg \phi'(x) = \theta$ then

$$|\phi'(x)| = |\phi'(e^{-i\theta}x)| = |\operatorname{Re} \phi'(e^{-i\theta}x)| = |\phi_r(e^{-i\theta}x)| \\ \le \|\phi\| \|e^{-i\theta}x\| = \|\phi\| \|x\|.$$

□

We denote by X^* the space of bounded linear functionals on X. It is a Banach space in its own right under the operator norm. (Completeness follows as in the proof of completeness of $B(H_1, H_2)$ on assignment 1.)

Corollary 14.17. *Let X be a normed space. Then*

(i) X^* is 'total', that is, it separates points;

(ii) $\|x\| = \sup\{|\phi(x)| : \phi \in X^*, \|\phi\| = 1\}$;

(iii) for Y a subspace of X and $d(x, Y) = d > 0$, there is $\phi \in X^*$ such that $\phi|_Y = 0, \|\phi\| = 1, \phi(x) = d$.

(iv) a point x lies in the closed linear span of $A \subset X$ if and only if $\phi(x) = 0$ for all $\phi \in X^*$ with $\phi|_A = 0$.

14.5 Examples of dual spaces

• Every Hilbert space is its own dual.

• The dual of $L^p(M, \mu)$. Observe that Hölder's inequality,

$$\Big| \int_M fg d\mu \Big| \le \|f\|_p \|g\|_q,$$

shows that every $g \in L^q$, $p^{-1} + q^{-1} = 1$, acts as a bounded linear functional on L^p by integration. Does every bounded linear functional arises in this way? You already know that the answer is yes when $p = 2$.

Theorem 14.18. *Let (M, μ) be σ-finite and suppose $1 < p < \infty$. Then the dual space of $L^p(M, \mu)$ is $L^q(M, \mu)$, where $p^{-1} + q^{-1} = 1$.*

Proof: Let us prove this first in the case where μ is finite. Let l be a bounded linear functional. Define a function $\nu : \mathcal{M} \to \mathbb{C}$ by

$$\nu(E) = l(1_E).$$

(Since μ is a finite measure, 1_E is in L^p for every measurable E.) I claim that ν is a complex measure. Clearly ν is finitely additive. To show countable additivity, suppose that E_i are disjoint measurable sets with union E. Let $F_n = \cup_{i=1}^N E_i$. Then

$$1_{F_n} \to 1_E \text{ in } L^p(M).$$

Hence, $l(1_{F_n}) \to l(1_E)$. It follows that $\nu(F_n) \to \nu(E)$, so ν is countably additive and hence a measure.

If E has measure zero, then 1_E is zero a.e. and hence $\nu(E) = 0$. It follows that ν is absolutely continuous with respect to μ. By the Radon-Nikodym theorem, there exists an integrable function g such that

$$l(1_E) = \nu(E) = \int_E g \, d\mu.$$

It follows that

$$l(f) = \int_E fg \, d\mu.$$

Finally we need to show that $g \in L^q(M, \mu)$. To do this, define $f_N = |g|^{q-1} e^{-i \arg g} 1_{|g| \le N}$. Then $f_N \in L^p$ and we have

$$l(f) = \int_{|g| \le N} |g|^q \, d\mu \le \|l\| \|f\|_p.$$

But

$$\begin{aligned} \|f\|_p &= \Big(\int_{|g| \le N} |g|^{(q-1)p} \, d\mu \Big)^{1/p} \\ &= \Big(\int_{|g| \le N} |g|^q \, d\mu \Big)^{1/p} \\ &= \Big(\int_{|g| \le N} |g|^q \, d\mu \Big)^{1-1/q} \end{aligned}$$

Rearranging, we get

$$\Big(\int_{|g| \le N} |g|^q \, d\mu \Big)^{1/q} \le \|l\|.$$

By the monotone convergence theorem we get

$$\|g\|_q = \lim_{N \to \infty} \Big(\int_{|g| \le N} |g|^q \, d\mu \Big)^{1/q} \le \|l\|,$$

showing that l is given by integration against some L^q function. Moreover, putting $f = |g|^{q-1} e^{-i \arg g}$ we see that $f \in L^p$ and $l(f) = \|g\|_q \|f\|_p$, so we have $\|l\| = \|g\|_q$.

In the case that μ is σ-finite, we take an increasing sequence of sets E_i of finite measure with union M, and apply the argument to l restricted to functions supported in E_i. We get a sequence of functions g_i supported in E_i, with $\|g_i\|_q \le \|l\|$. It is not hard to show that g_i and g_j agree a.e. on $E_i \cap E_j$ and so this determines a function $g \in L^q$ on M (a.e.) that implements l. □

Example. Suppose that $f \in L^q(\mathbb{R}^n)$ and $k \in L^r(\mathbb{R}^n)$. Let $q^{-1} + r^{-1} = 1 + p^{-1}$. Then $f * k \in L^p(\mathbb{R}^n)$, and

$$\|f * k\|_p \le \|f\|_q \|k\|_r.$$

To show this, in view of the theorem above, it suffices to show that $f * k$ is a bounded linear functional on $L^s(\mathbb{R}^n)$, where $s^{-1} + p^{-1} = 1$, with norm at most $\|f\|_q \|k\|_r$. So integrate against a function $g \in L^s$: we need to show that

$$\int g(x) f(x-y) k(y) \, dx \, dy \le \|f\|_q \|k\|_r \|g\|_s. \tag{14.2}$$

To do this, we need the generalized Hölder's inequality:

$$\int h_1 h_2 h_3 \le \|h_1\|_{a_1} \|h_2\|_{a_2} \|h_3\|_{a_3}$$

provided that $a_1^{-1} + a_2^{-1} + a_3^{-1} = 1$. Now, since $q^{-1} + r^{-1} + s^{-1} = 2$, we can find a_i as above satisfying

$$\begin{aligned} \frac{1}{q} &= \frac{1}{a_1} + \frac{1}{a_3} \\ \frac{1}{r} &= \frac{1}{a_2} + \frac{1}{a_3} \\ \frac{1}{s} &= \frac{1}{a_1} + \frac{1}{a_2} \end{aligned}$$

Now write the LHS of (14.2) in the form

$$\int (g^{\frac{a_2}{a_1+a_2}} f^{\frac{a_3}{a_1+a_3}})(g^{\frac{a_1}{a_1+a_2}} k^{\frac{a_3}{a_2+a_3}})(k^{\frac{a_2}{a_2+a_3}} f^{\frac{a_1}{a_1+a_3}})$$

and apply generalized Hölder with exponents a_i. We get,

$$\Big(\int |g(x)|^s |f(x-y)|^q\,dx\,dy\Big)^{1/a_1}\times$$
$$\Big(\int |g(x)|^s |k(y)|^r\,dx\,dy\Big)^{1/a_2}\times$$
$$\Big(\int |f(x-y)|^q |k(y)|^r\,dx\,dy\Big)^{1/a_3}$$
$$= \|f\|_q\|k\|_r\|g\|_s$$

as required.

Another important case is $X = C(S)$, the continuous functions on a compact metric space S (or more generally, a compact Hausdorff topological space S) with the sup norm.

Theorem 14.19 (Riesz representation theorem)**.** *The dual space of $C(S)$ is $M(S,\mathcal{B})$, the space of complex Borel measures on S.*

• $M(S) = M(S,\mathcal{B})$ acts on $C(S)$ by integration: if ν is a complex Borel measure, then any $f \in C(S)$ is integrable w.r.t. ν and

$$\int_S f\,d\nu \le \|f\|_\infty \|\nu\|.$$

According to the theorem, every bounded linear functional on $C(S)$ is of this form.

The proof of this theorem is fairly long. We'll break these into lemmas (each of which has a somewhat technical, not but difficult, proof). We first note that it is enough to treat real linear functionals (those that take real values on real functions), since every linear functional is of the form $J_1 + iJ_2$ where J_i are real linear functionals.

The first is about positive linear functionals. We say that $l \in C(S)^*$ is positive if $l(f) \ge 0$ whenever $f \ge 0$.

Lemma 14.20. *Every real bounded linear functional F on $C(S)$ can be written as the difference of two positive linear functionals $F_+ - F_-$, with the property that $\|F\| = \|F_+\| + \|F_-\|$.*

Given this result, it is enough to show that every positive linear functional can be represented by a finite Borel (positive) measure.

Lemma 14.21. *Every positive linear functional J on $C(S)$ determines a metric exterior measure μ^* according to the formula*

$$\mu^*(E) = \inf_O \sup_f J(f)$$

where we sup over all $f \in C(S)$ such that $0 \le f \le 1$ with support in O, and then inf over all open sets O containing E.

By a theorem from lectures several weeks ago (SS Theorem 1.2 chapter 6), μ^* determines a unique Borel measure μ.

Lemma 14.22. *The measure μ implements J and is such that the total variation of μ is equal to $\|J\|$.*

The proofs of these Lemmas were omitted in class, but are included here if you are interested in reading them.

Proof of Lemma 14.21: It is obvious that $\mu^*(\emptyset) = 0$ and that $E_1 \subset E_2$ implies that $\mu^*(E_1) \le \mu^*(E_2)$. It remains to show countable subadditivity. To streamline the notation, we write '$f \sim O$' to mean '$f \in C(S)$, $\operatorname{supp} f \subset O$ and $0 \le f \le 1$'.

First we show that $\mu^*(E_1 \cup E_2) \le \mu^*(E_1) + \mu^*(E_2)$. It is enough to show that $\mu^*(E_1 \cup E_2) \le \mu^*(E_1) + \mu^*(E_2) + 2\epsilon$ for arbitrary $\epsilon > 0$. To do this, we need to find some open set O containing $E = E_1 \cup E_2$, such that $J(f) \le \mu^*(E_1) + \mu^*(E_2) + 2\epsilon$ for all $f \sim O$. Choose $O_i \supset E_i$ such that

$$\sup_{f\sim O_i} J(f) \le \mu^*(E_i) + \epsilon.$$

Let $O = O_1 \cup O_2$ and let $f \sim O$. It is enough to show that f can be decomposed $f = f_1 + f_2$, where $f_i \sim O_i$. To do this we use Urysohn's Lemma which says that if C_1, C_2 are two disjoint closed subsets of a compact metric space, then there exists a continuous function g such that $0 \le g \le 1$, $g = 0$ on C_1, $g = 1$ on C_2. We apply this with $C_1 = (\operatorname{supp} f \cap O_1^c)_{\delta/3}$ and $C_2 = (\operatorname{supp} f \cap O_2^c)_{\delta/3}$ where $\delta > 0$ is the distance between the disjoint closed subsets $\operatorname{supp} f \cap O_i^c$, and $(\cdot)_{\delta/3}$ is the $\delta/3$ enlargement. Then $f_2 = f(1-g)$ and $f_1 = fg$ is such a decomposition, establishing that μ^* is finitely subadditive.

To check countable subadditivity, consider a sequence $E_1, E_2, \dots$ of Borel subsets with union E. It is sufficient to check, in view of the previous result, that $\mu^*(E) \le \lim_{n\to\infty} \mu^*(\cup_{i=1}^n E_i)$. Wlog, suppose that E_n is an increasing sequence; then we need to show that $\mu^*(E) \le \lim_{n\to\infty} \mu^*(E_n)$. So choose, for given $\epsilon > 0$, an open set $O_n \supset E_n$ such that

$$\sup_{f\sim O_n} J(f) \le \mu^*(E_n) + \epsilon. \tag{14.3}$$

Moreover, we can do this so that the O_n are increasing. To do this we first choose $O_j' \supset E_j \setminus E_{j-1}$ so that (14.3) is satisfied with E_j replaced by $E_j \setminus E_{j-1}$ and ϵ replaced by $\epsilon 2^{-j}$, and then take $O_n = \cup_{j=1}^n O_j'$.

Now take $O = \cup O_i$. Consider any $f \sim O$. Then since O_n is an open cover of $\operatorname{supp} f$ which is compact, some finite number of the O_n cover $\operatorname{supp} f$. Since the O_n are increasing, in fact one O_N covers $\operatorname{supp} f$. But then, we have $J(f) \le \mu^*(E_N) + \epsilon \le \lim_{n\to\infty} \mu^*(E_n) + \epsilon$. Taking the sup over $f \sim O$ gives $\mu^*(E) \le \lim_{n\to\infty} \mu^*(E_n) + \epsilon$, and since ϵ is arbitrary we finally get $\mu^*(E) \le \lim_{n\to\infty} \mu^*(E_n)$. Hence μ^* is countably subadditive.

Finally we show that μ^* is a metric exterior measure. This means that we just show that $\mu^*(E_1) + \mu^*(E_2) = \mu^*(E_1 \cup E_2)$ whenever $d(E_1, E_2) > 0$. To do this we consider disjoint open sets O_i with $O_i \supset E_i$. Then, for any set $O \supset (E_1 \cup E_2)$, we have

$$\begin{aligned}\sup_{f\sim O} J(f) &\ge \sup_{f\sim O\cap(O_1\cup O_2)} J(f)\\ &= \sup_{f_1\sim O\cap O_1} J(f_1) + \sup_{f_2\sim O\cap O_2} J(f_2)\\ &\ge \mu^*(E_1) + \mu^*(E_2).\end{aligned}$$

The equality in the second line is because an $f \sim O_1 \cup O_2$ is precisely a pair of $f_i \sim O_i$. Taking the inf over all such open sets O, we obtain $\mu^*(E_1 \cup E_2) \ge \mu^*(E_1) + \mu^*(E_2)$, and the opposite inequality is just subadditivity. This establishes that μ^* is a metric exterior measure and completes the proof. □

Proof of Lemma 14.22: First, consider a nonnegative function f. We will show that

$$J(f) = \int_X f\,d\mu.$$

This is clear for the function $f \equiv 1$, since $\int_X 1\,d\mu = \mu(X) = \mu^*(X) = \sup_{0\le f\le 1} J(f) = J(1)$ by the positivity of J. Now consider an arbitrary nonnegative continuous function f. We show that $\int f \le J(f)$. To do this, we define the simple function

$$f_n = 2^{-n}\sum_{j=0}^{\infty} 1_{E_{j,n}}, \quad E_{j,n} = \{x \in S \mid f(x) < (j+1)2^{-n}\}.$$

This is a nonnegative simple function with $f_n \le f$ and $f_n \uparrow f$. Therefore,

$$\int f\,d\mu = \lim_n \int f_n\,d\mu.$$

But

$$\begin{aligned}\int f_n\,d\mu &= 2^{-n}\sum_j \mu(E_{j,n})\\ &= 2^{-n}\sum_j \sup_{g_{j,n}\sim E_{j,n}} J(g_{j,n})\\ &= \sup_{g_{j,n}\sim E_{j,n}} J\Big(\sum_j 2^{-n} g_{j,n}\Big)\\ &\le J(f)\end{aligned}$$

since the function $\sum_j 2^{-n} g_{j,n}$ is a continuous function that is $\le f$ pointwise. (Also note that we used the fact that $E_{j,n}$ are open to eliminate the inf over open sets.) Thus we have proved that $\int f \le J(f)$. But also, choosing M sufficiently large that $M \ge f$, we have $\int(M-f) \le J(M-f)$. Since $\int M = J(M)$ as observed above, we deduce that $\int(-f) \le J(-f)$. Therefore, $\int f = J(f)$ for all nonnegative functions f. By linearity, this is true for all complex-valued f.

To prove that the $\|\mu\|_{M(S)} = \|J\|$, we note that $\|\mu\|_{M(S)} = \int 1\,d\mu = J(1) = \|J\|$ using the positivity of J. This completes the proof. □

Proof of Lemma 14.20: Let F be a real functional on $C(S)$. Define, for $f \ge 0 \in C(S)$,

$$F_+(f) = \sup_{0\le g\le f} F(g).$$

Then $F_+(f) \ge F(f)$, $F_+(f) \ge 0$.

It is true, but not obvious, that F_+ is linear. To show this, first note $F_+(cf) = cF_+(f)$ $(f \ge 0)$. If $f_i \ge 0$, then $F_+(f_1 + f_2) \ge F_+(f_1) + F_+(f_2)$ since if $0 \le g_i \le f_i$ then $0 \le g_1 + g_2 \le f_1 + f_2$. On the other hand, if $0 \le \psi \le f_1 + f_2$, then $\psi = \psi_1 + \psi_2$ where $\psi_1 = \min(f_1, \psi)$ and $\psi_2 = \psi - \min(f_1, \psi)$ are between 0 and f_1, resp. f_2, which gives us $F_+(f_1 + f_2) \le F_+(f_1) + F_+(f_2)$, so F_+ is linear on nonnegative functions.

To extend to all real functions, we define $F_+(f) = F_+(M+f) - F_+(M)$ for sufficiently large M. This is well-defined since

$$\begin{aligned}&F_+(M+f) + F_+(N) = F_+(N+f) + F_+(M)\\ \Longrightarrow\ &F_+(M+f) - F_+(M) = F_+(N+f) - F_+(N)\end{aligned}$$

for sufficiently large M, N. It is not hard to check that this function is linear. Then $F_- = F_+ - F$ is also linear and nonnegative.

Finally, to show that $\|F\| = \|F_+\| + \|F_-\|$, it suffices to show that $\|F\| \ge \|F_+\| + \|F_-\|$ since the reverse inequality is just the triangle inequality. To prove this, take any function g with $0 \le g \le 1$, and compute

$$\|F\| \geq F(2g-1) = 2F(g) - F(1).$$

Taking the sup over all g, we find that

$$\|F\| \geq 2F_+(1) - F(1) = F_+(1) + F_-(1) = \|F_+\| + \|F_-\|.$$

□

14.6 Weak topologies

To motivate the study of weak topologies, let us begin with

Theorem 14.23. *A normed space X has a compact unit ball if and only if X is finite dimensional.*

To prove this we need the Riesz lemma:

Lemma 14.24 (Riesz)**.** *Let M be a proper closed subspace of a normed space X. Then for each $0 < \alpha < 1$ there is $x \in X$, $\|x\| = 1$ such that $d(x, M) > \alpha$.*

Proof: Take $y \in X \backslash M$, so that $d(y, M) = \delta > 0$. Since $\alpha < 1$ there is $z \in M$ with $\|z - y\| < \delta/\alpha$. Then $x = (z - y)/\|z - y\|$ suffices. □

Proof of Theorem 14.23: If X has finite dimension the result is clear. Otherwise, define a (linearly independent) sequence (x_n) as follows. Having chosen $x_1, \ldots, x_n$, let $M_n = \operatorname{span}\{x_1, \ldots, x_n\}$ and by Lemma 14.24 take $x_{n+1} \in X \setminus M_n$ of norm one and such that $d(x_{n+1}, M_n) \geq 1/2$. Then (x_n) has no convergent subsequence. □

This makes life difficult for solving analytic problems, because a very common strategy for finding solutions to PDEs or other problems where we are looking for a function is to obtain the solution as the limit of approximate solutions. In some good cases we can obtain the limit from a convergent sequence, but in many cases we need to use compactness and extract a convergent subsequence. The above result says that the norm topology is no good for this, when we are working in an infinite-dimensional Banach space (the usual situation). We need to consider weaker topologies.

What do weaker topologies gain for us? These have *fewer open sets*, and hence there are *fewer open covers* of a given set K, and hence it is more likely that we can find a finite subcover, and hence there are *more compact sets*. Moreover there are more convergent sequences! To pay for all this, we find that fewer functions are continuous, when we change to a weaker topology.

Where do weaker topologies come from? From dual spaces! We define the $\sigma(X, X^*)$-topology on X to be the weakest topology that makes all the linear functionals $\phi \in X^*$ continuous. This is a much weaker topology than the norm topology.

We say $x_n \xrightarrow{\text{weakly}} x$ to mean $\phi(x_n) \to \phi(x)$ for all $\phi \in X^*$. (Unfortunately, because the weak topology is not metrizable, this does not completely characterize it.) The sets $\phi^{-1}(U)$, where $\phi \in X^*$ and $U \subset \mathbb{C}$ is open, form a sub-base for the topology. This means that $V \subset X$ is open in the $\sigma(X, X^*)$ topology precisely if it is a (possibly infinite) union of finite intersections of sets of the form $\phi^{-1}(U)$.

Unfortunately we do not have time for a detailed study of the $\sigma(X, X^*)$ topology. The following exercises outline many of the important properties. You can find most of these, and more, covered in the notes at `http://perso.crans.org/lecomte/Math/WeakTopologies.pdf`.

Exercise. Show that if X is infinite dimensional, all open sets in the $\sigma(X, X^*)$-topology are unbounded.

Exercise. The weak topology is Hausdorff (that is, given $x \neq y$, we can find non-intersecting open sets $U \ni x$ and $V \ni y$), and hence limits of sequences are unique. (Hint: use Hahn-Banach.)

Exercise. Show that in an infinite-dimensional Hilbert space, an orthonormal basis $e_1, e_2, \ldots$ converges to 0 weakly.

Exercise. Show that a sequence (x_n) in a Banach space X converges to x weakly iff $x^*(x_n)$ converges to $x^*(x)$ for every $x^* \in X^*$.

Exercise. Show that a sequence (x_n) in a Banach space X converges to x weakly iff $x^*(x_n)$ converges to $x^*(x)$ for every x^* in a dense subset of the unit ball in X^*.

Exercise. Show that the closed unit ball in X is also closed in the weak topology.

Exercise. Show that a bounded linear transformation $T : X \to Y$ is continuous if both X and Y are given the weak topology.

Exercise. The weak topology and the norm topology coincide iff X is finite dimensional.

Is the closed unit ball compact in this topology? Sometimes! For example, it is true for Hilbert spaces. But it is not true for $L^1(\mathbb{R})$. There is a closely related topology that *is* compact for the unit ball; however, it only applies to dual spaces.

Exercise. Find a sequence in the unit ball of $L^1([0, 1])$ that has no convergent subsequence in the weak topology.

14.7 The second dual, and the weak−∗ topology

Since the dual X^* of a Banach space X is itself a Banach space, we can form its dual space $(X^*)^*$, which we write X^{**} and call the second dual or the bidual space of X. In fact, X^{**} is somewhat more closely related to X than X^* is. To see this, note that any element x of X acts as a bounded linear functional on X^*, by defining

$$x(x^*) = x^*(x), \qquad x \in X,\ x^* \in X^*.$$

Notice that $|x(x^*)| \leq \|x\|\|x^*\|$. So the norm of x as an element of X^{**} is at most its norm in X. By the Hahn-Banach theorem, we can choose for any $x \in X$ an $x^* \in X^*$ such that $|x(x^*)| = \|x\|\|x^*\|$, which implies that the norm of x as an element of X^{**} is precisely the norm of x in X. Therefore, there is a continuous injection $J : X \to X^{**}$ which is an isometry onto its image (which is therefore a closed subspace of X^{**}).

It is now irresistible to ask whether $J(X) = X^{**}$. The answer is sometimes. For example, it is true for any Hilbert space. But it is not true if $X = C([0, 1])$. To see this, recall that the dual space of $C([0, 1])$ is the space of bounded Borel measures on $[0, 1]$ with the total variation norm. It is not hard to see that the function $1_{[0,1/2]}$ acts as a bounded linear functional on this space. Indeed, every bounded Borel-measurable function is included in the second dual of $C([0, 1])$. Another example is the usual suspect $L^1(\mathbb{R})$; the second dual of this space is strictly larger than itself.

We call X *reflexive* if $J(X) = X^{**}$.

On the dual space X^* of a Banach space X, we have the weak topology $\sigma(X^*, X^{**})$. But we can define another topology, namely $\sigma(X^*, X)$, called the weak-∗ topology that is the weakest topology such that all the linear functionals $x \in J(X) \subset X^{**}$ are continuous. This is weaker (that is, not stronger) than the weak topology (i.e. fewer open sets, more convergent sequences, more compacts sets, fewer continuous functions). We shall shortly prove the very important result that the closed unit ball in X^* is compact in the weak-∗ topology.

Exercise. Show that the weak-∗ topology is Hausdorff. (No Hahn-Banach theorem required: if $\phi \neq \psi$, there is some x and r so $\phi(x) < r < \psi(x)$. Use this to write down disjoint weak-∗ open neighbourhoods.)

14.8 Banach-Alaoglu theorem

There are various versions of the Banach-Alaoglu theorem.

Theorem 14.25 (Banach-Alaoglu, version 0)**.** *Suppose that X is separable. Then the unit ball of X^* is sequentially compact in the weak-∗ topology.*

Theorem 14.26 (Banach-Alaoglu, version 1)**.** *Suppose that X is separable. Then the unit ball of X^* is compact in the weak-∗ topology.*

Theorem 14.27 (Banach-Alaoglu, general version)**.** *The unit ball of X^* is compact in the weak-∗ topology.*

Corollary 14.28. *If X is reflexive, the closed unit ball of X is weakly compact.*

(Kakutani's theorem states the converse is also true.)

Note that separability is required for sequential compactness: ℓ^∞ is not separable, and the sequence of functionals $a \mapsto a(n)$ does not have a weak-∗ convergent subsequence.

To prove these, we'll need the Tychonoff theorem, asserting that a product of compact spaces is compact. For versions 0 and 1 above, we only need a weaker version of the Tychonoff theorem.

Theorem 14.29 (Tychonoff – metric case)**.** *Let (X_n, d_n) be a sequence of compact metric spaces. Then $\prod X_n$ is compact.*

Proof: Define a metric on $X = \prod X_n$ by

$$d((x_n), (y_n)) = \sum_{n=1}^{\infty} 2^{-n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}.$$

Check that this gives the product topology on X, and that convergence in this topology is the same as convergence of every component.

It suffices to show any sequence in X has a convergent subsequence. Let $(x^{(m)})$ be a sequence in X, here $x^{(m)} = (x_n^{(m)})_{n\geq 1}$. Then by compactness of X_1, there is a subsequence $x^{(m,1)}$ of $x^{(m)}$ such that $(x_1^{(m,1)})$ converges in X_1. Then there is a subsequence $x^{(m,2)}$ of $x^{(m,1)}$ such that $(x_2^{(m,2)})$ converges in X_2, and so on. Then $(x_n^{(m,m)})$ converges in X_n for every $n \in \mathbb{N}$. But this means that $(x^{(m,m)})$ converges in X. □

Before stating the general version of the Tychonoff theorem, recall that the product topology on $\prod X_\alpha$ is the weak topology determined by the projection maps $\pi_\beta : \prod X_\alpha \to X_\beta$, so basic open sets are of the form $\prod U_\alpha$ where each $U_\alpha \subseteq X_\alpha$ is open, and for all but finitely many α, $U_\alpha = X_\alpha$.

Theorem 14.30 (Tychonoff)**.** *The product of a family of compact topological spaces is compact.*

Definition 14.31. A collection of sets has the *finite intersection property* (FIP) if every finite subcollection has non-empty intersection.

Recall that a set is compact if and only every every collection of *closed* sets with the FIP itself has non-empty intersection.

Proof: Let $X = \prod X_\alpha$, where each X_α is compact. Let F be a family of closed subsets of X with the finite intersection property. A Zorn's lemma argument gives a maximal family H containing F and having the FIP (though sets in H are not necessarily closed). Note that by maximality H is necessarily closed under finite intersections, and any set which meets every set in H must also be in H.

For a given index α, $\{\pi_\alpha(A) : A \in H\}$ has the FIP in X_α. By compactness, the closures of these sets must thus have non-empty intersection, let $f(\alpha)$ be a point in the intersection. Then $f = (f(\alpha)) \in X$. For an index β, neighbourhood U_β of $f(\beta)$, and $A \in H$, $U_\beta \cap \pi_\beta(A) \neq \emptyset$, that is, $\pi_\beta^{-1}(U_\beta) \cap A \neq \emptyset$. But then $\pi_\beta^{-1}(U_\beta) \in H$. But then any finite intersection of such sets also lies in H, and so must meet every element of H since H has the FIP. These sets, the finite intersections of sets $\pi_\beta^{-1}(U_\beta)$, give a neighbourhood base for f: that is, given any open set containing f, no matter how small, there is one of these sets properly contained in it, still containing f. It follows that f lies in the closure of every element of H, and so in particular lies in each member of F. $\square$

Proof of Banach-Alaoglu, version 0: Take a sequence x_n in X which is dense in the unit ball. Given a sequence ϕ_j in $\overline{B(0,1)} \subset X^*$, we extract a subsequence ϕ_j^1 such that $\phi_j^1(x_1)$ converges. Then we extract a subsequence ϕ_j^2 of this subsequence such that $\phi_j^2(x_2)$ converges, and so on. The diagonal subsequence ϕ_j^j has the property that $\phi_j^j(x_n)$ converges for all n. (This is because ϕ_j^j is a eventually a subsequence of ϕ_j^n.) We then need to show that $\phi_j^j(x)$ converges for all $x \in B(0,1) \subset X$. Take $\{y_m\} \subset \{x_n\}$ converging to $x \in B(0,1)$. We have $\lim_j \phi_j^j(y_m) = z_m$ for some $z_m \in \mathbb{C}$, and moreover $\lim_m z_m = z$ for some $z \in \mathbb{C}$ (since $|\phi_j^j(y_m - y_{m'})| \leq \|y_m - y_{m'}\|$). Now we show $\lim_j \phi_j^j(x) = z$. Choose m large enough so $\|y_m - x\| < \epsilon$ and $|z_m - z| < \epsilon$. Then choose j large enough so $|\phi_j^j(y_m) - z_m| < \epsilon$. Then we have $|\phi_j^j(x) - z| < 3\epsilon$. Now define $\phi(x)$ to be $\lim_j \phi_j^j(x)$. Since this is given as a pointwise limit of a sequence of bounded linear functionals of norm at most 1, ϕ is linear with $\|\phi\| \leq 1$. Then $\phi_j^j \to \phi$ in the weak-$*$ topology. $\square$

Proof of Banach-Alaoglu, version 1: In this case, we take a countable dense subset $\{x_n\}$ of the unit ball in X, and define a product space $D = \prod_n D_{x_n}$, where

$$D_x = \{z \in \mathbb{C} : |z| \leq \|x\|\}.$$

Then D is a compact metric space with the metric as defined in the proof of the metric space Tychonoff theorem. The metric space topology on D coincides with the product topology. In this

case, it is easier to think of it as generated by basis sets in the product topology, that is by finite intersections of sets of the form $\pi_{x_n}^{-1}(G)$ where G is open in $\mathbb{C}$.

Define a map F from $\overline{B(0,1)} \subset X^*$ to D by

$$\phi \mapsto (\phi(x_n))_{n=1}^\infty.$$

I claim that F is a homeomorphism onto its image. To check this, we need to show that F is continuous, one-to-one, and open (so that the inverse is also continuous). Injectivity follows from the density of the sequence x_n. To prove continuity, consider an basis open set in D. It is a finite intersection of sets of the form $\pi_{x_n}^{-1}(G)$ where G is open in $\mathbb{C}$. The inverse image (under F) of such sets are open for the weak-$*$ topology on $\overline{B(0,1)}$ so F is continuous.

We still need to show that F is an open map, i.e. maps open sets to open sets. So consider an open set in $\overline{B(0,1)}$; it suffices to consider sets of the form $U_{x,G} = \{\phi \mid \phi(x) \in G\}$ for a fixed $x \in X$ and $G \subset \mathbb{C}$ open.

Choose $\phi_0 \in U_{x,G}$, and let $\epsilon = d(\phi_0(x), \partial G) > 0$. We can find N such that $\|x_N - x\| < \epsilon/2$. Now consider the set $\{\phi \mid |\phi(x_N) - \phi_0(x_N)| < \epsilon/2\}$. This is an open set contained in the set $U_{x,G}$, and the image under F of this set is open. Therefore, F is an open map. Therefore, F is a homeomorphism onto its image.

Finally, we show that range$F$ is closed in D. To do this, suppose that $F(\phi_j)$ converges to a point $(y_n) \in D$ as $j \to \infty$. Then, it is not hard to show that $\phi_j(x)$ converges to some number $\phi(x)$ for every $x \in B(0,1) \subset X$. (Use the fact that $|\phi_j(x_1) - \phi_j(x_2)| \leq \|x_1 - x_2\|$.) It is straightforward to check that ϕ is linear, and $\|\phi\| \leq 1$. Therefore, $(y_n) = F(\phi)$, and we see that range$F$ is closed. But since D is compact, also range$F$ is compact, and hence $\overline{B(0,1)} \subset X^*$, which is homeomorphic to range$F$, is compact (in the weak-$*$ topology). $\square$

Proof of Banach-Alaoglu, general version: For $x \in X$, set $D_x = \{z \in \mathbb{C} : |z| \leq \|x\|\}$. Then $D = \prod D_x$ is compact. Define the injective map $F : \overline{B(0,1)} \subset X^* \to D$ by $F(\phi) = (\phi(x))_{x \in X}$. Consider the topology induced on $\overline{B(0,1)}$ by F. Open sets are given by finite intersections of sets of the form

$$\pi_x^{-1}(G), \quad G \subset \mathbb{C} \text{ open}.$$

Here x is an arbitrary fixed element of X. This corresponds on $\overline{B(0,1)}$ to finite intersections of sets of the form

$$\{\phi \mid \phi(x) \in G\}.$$

But this is precisely the weak*-topology on $\overline{B(0,1)}$. Thus, with this topology on $\overline{B(0,1)}$, the map F is a homeomorphism onto its image. Its image consists of the 'linear' elements of D. To check

that the image is closed, take a point $(y_x)_{x \in X}$ in D that is not in the image of F. Then this point is not 'linear', so there must be a linear relation $x = ax_1 + bx_2$ in X ($a, b \in \mathbb{C}$) with $y_x \neq ay_{x_1} + by_{x_2}$. Let $\epsilon = |y_x - ay_{x_1} - by_{x_2}|$. Now consider the open set U of points $(z_x)_{x \in X}$ such that

$$|z_x - y_x| < \frac{\epsilon}{3}, \;\; |a||z_{x_1} - y_{x_1}| < \frac{\epsilon}{3}, \;\; |b||z_{x_2} - y_{x_2}| < \frac{\epsilon}{3}.$$

Then U is disjoint from the image of F. Thus, the image of F is closed, and therefore compact. Since $\overline{B(0,1)}$ with the weak topology is homeomorphic to the image of F, it is also compact. $\square$

14.9 Applications of Banach-Alaoglu

14.10 Minimizing functionals on a Banach space

Let X be a Banach space (for simplicity, we shall consider only real Banach spaces from now on). Consider the problem of minimizing some nonlinear function $F : X \to \mathbb{R}$, which we assume to be bounded below.

In general, this will not be possible. This can be for many different reasons. Here are some examples that illustrate different ways in which minimizers fail to exist:

- $X = \mathbb{R}$, $F(x) = e^x$.
- $X = \mathbb{R}$, $F(x) = |x|$ for $x \neq 0$, $F(0) = a > 0$.
- $X = L^2([0,1])$, and

$$F(f) = \int_0^1 x f(x)^2 \, dx + (\|f\|_2 - 1)^2.$$

It is easy to see that $\inf F = 0$. However, there is no function $f \in L^2([0,1])$ such that $F(f) = 0$. We will need some conditions on F. What should they be?

Recall the theorem that says that a continuous function on a compact space always attains its minimum. If X is reflexive, e.g. $X = L^p(M, \mu)$ for $1 < p < \infty$), which we shall assume from now on, we have just seen that the closed unit ball on X is compact for the weak topology. So if F is continuous in the weak topology, and if $F(x) \to \infty$ as $\|x\| \to \infty$, we would know that F attains its minimum.

This is true, and sounds good, but the class of functions continuous in the weak topology is too small to be of genuine interest.

Example. The function $x \mapsto \|x\|$ on a Hilbert space H is not weakly continuous. For if it were, then the set $S = \{x \mid \|x\| = 1\}$ would be a weakly closed set. But we can find a sequence

of elements of S, such as an orthonormal sequence, that converges weakly to zero. So S is not weakly closed.

However, in order to *minimize* functionals, continuity is an unnecessarily restrictive condition: lower semicontinuity is enough. We say that a function F on a topological space is (sequentially) lower semicontinuous iff, for any sequence $x_n \to x$, we have $F(x) \leq \liminf F(x_n)$. (The correct definition on lower semicontinuity should be that for any net $x_\alpha \to x$, we have $F(x) \leq \liminf F(x_\alpha)$. In fact, the difference between sequential lower semicontinuity and lower semicontinuity won't be important in what follows, so we'll ignore the difference.)

For example, the function $g(t) : \mathbb{R} \to \mathbb{R}$ which is 0 for $t \leq 0$ and 1 for $t > 0$ is lower semicontinuous. The function F in the second example above is upper semicontinuous; if $a < 0$ it would be lower semicontinuous and then there is no problem finding a minimizer.

Lemma 14.32. *Let Z be a sequentially compact topological space and F a lower semicontinuous real function. Then F is bounded below and attains its minimum.*

Proof: Take a minimizing sequence, i.e. a sequence of points x_n such that $F(x_n) \to \inf F$ (tending to $-\infty$ if $\inf F = -\infty$). Let x be a limit point of a subsequence of the x_n, which we continue to denote (x_n). Then $F(x) \leq \liminf F(x_n) = \inf F$. So the inf is finite, and attained at any limit point of the x_n. $\square$

Lemma 14.33. *Let Z be a compact topological space and F a lower semicontinuous real function. Then F is bounded below and attains its minimum.*

Proof: First, suppose for contradiction that $\inf F = -\infty$. Define closed sets $C_n = \{x | F(x) \leq -n\}$. These sets have the finite intersection property, and since Z is compact the complete intersection is non-empty, which is impossible.

Now that we know $\inf F$ is finite, define closed sets $D_n = \{x | F(x) \leq \inf F + 1/n\}$. Again, these sets have the finite intersection property and since Z is compact the complete intersection is non-empty. Any point in this complete intersection realizes the infimum. $\square$

So now the question is, which functions on X are lower semicontinuous? It turns out that convex functions (with a couple of other additional assumptions) have this property. We say that a function F mapping a real vector space V to $(-\infty, \infty]$ is convex if for every $x, y \in V$ and $0 \leq \lambda \leq 1$ we have

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y).$$

Note that we allow $+\infty$ as a value of F here.

For example, the function

$$F(u) = \int_{\mathbb{R}^n} u(x)^2 + u(x)^4\, dx$$

is convex on the (real) Banach space $L^2(\mathbb{R}^n)$. This follows from the fact that $t \mapsto t^2 + t^4$ is convex on $\mathbb{R}$.

Theorem 14.34. *Suppose that F is a convex function on a Banach space X, such that F is norm lower semicontinuous. Then F is weakly lower semicontinuous.*

To prove this, we first need the following lemma about convex functions:

Lemma 14.35. *Let F be a convex function on a real vector space X. Then for all $x \in X$ such that $F(x) < \infty$, there is an affine function ϕ (i.e. linear plus constant) such that $\phi(x) = F(x)$ and $\phi \leq F$ on X.*

The proof is almost identical to version 1 of the Hahn-Banach theorem.

Proof: First, translate x to the origin and add a constant to F so that $F(0) = 0$. We now look for a linear function which is $\leq F$. Suppose that we have found a linear function on a subspace Y satisfying the conditions. We first show that if Y' is the subspace spanned by Y and a single element $z_0 \in X \setminus Y$, then we can extend ϕ to Y' such that $\phi \leq F$ on Y'. It is convenient to subtract the linear function $z + tz_0 \to \phi(z)$ (defined on Y') from both F and ϕ. Thus, we may assume wlog that $\phi = 0$ and $F \geq 0$ on Y.

Then since F is convex, for any $z', z'' \in Y$, and any $t_1, t_2 > 0$, we have

$$t_2 F(z'' + t_1 z_0) + t_1 F(z' - t_2 z_0) \geq 0.$$

Therefore, for arbitrary $z', z'' \in Y$ we have

$$\frac{F(z'' + t_1 z_0)}{t_1} \geq -\frac{F(z' - t_2 z_0)}{t_2}. \tag{14.4}$$

Define β to be any finite number between the inf of the LHS and the sup of the RHS in the above expression. Now define $\phi(z + tz_0) = t\beta$. Then for any $t > 0$, we have

$$\phi(z + tz_0) = t\beta \leq t\frac{F(z + tz_0)}{t} = F(z + tz_0).$$

On the other hand, for $t > 0$ we have

$$\phi(z - tz_0) = -t\beta \leq t\frac{F(z - tz_0)}{t} = F(z - tz_0).$$

Thus we have extended ϕ from Y to Y' so as to be $\leq F$.

The result now follows by Zorn's Lemma as in the proof of Hahn-Banach. □

Corollary 14.36 (Separating Hyperplane theorem)**.** *Let C be a closed convex set in a real Banach space X, and x a point in $X \setminus C$. Then there is a continuous affine functional $\phi : X \to \mathbb{R}$ such that $\phi \leq 1$ on C, and $\phi(x) > 1$.*

Proof: First translate C so that it contains the origin. Then we replace it by its ϵ-enlargement C_ϵ, where

$$C_\epsilon = \{z \in X \mid \exists c \in C \text{ such that } \|z - c\| \leq \epsilon\}$$

for $0 < \epsilon < d(C, x)$. Then C_ϵ is also closed and convex and contains a ball of radius ϵ around the origin. Define $\phi(y) = \sup\{\lambda \geq 0 \mid y \notin \lambda C_\epsilon\}$ for $y \neq 0$, and $\phi(0) = 0$. Then ϕ is convex and bounded by ϵ^{-1} times the norm. Also, $\phi \leq 1$ on C, and $\phi(x) > 1$. The result then follows from the previous lemma. □

Proof of Theorem: Consider a sequence x_n in X converging weakly to x. We must show that $F(x) \leq \liminf F(x_n)$. Let $a = \liminf F(x_n)$. It suffices to show that $F(x) \leq a + \epsilon$ for every $\epsilon > 0$.

By passing to a subsequence, we may assume that $F(x_n)$ converges to a, and hence that $F(x_n) \leq a + \epsilon$ for all n large enough. Consider the set $C = \{z \in X \mid F(z) \leq a + \epsilon\}$. This is a convex set, and is norm closed due to the norm lower semicontinuity of F. I claim that the limit point x lies in C. For if not, then by the Separating Hyperplane theorem, there exists an affine linear functional such that $\phi \leq 1$ on C and $\phi(x) > 1$. But this is a contradiction, since weak convergence requires that $\phi(x_n) \to \phi(x)$. Therefore, x lies in C. That means that $F(x) \leq a + \epsilon$ as claimed. This finishes the proof. □

Putting our results together, we have

Theorem 14.37. *Suppose that F is a convex function on a reflexive Banach space X, such that $F(x) \to \infty$ as $|x| \to \infty$, and such that F is norm lower semicontinuous. Then F attains its minimum value. Moreover, if F is strictly convex, then the minimum point is unique.*

Proof: We take a minimizing sequence as before. Since $F \to \infty$ as $|x| \to \infty$, this minimizing sequence lies in a fixed closed ball in X, which is weakly compact by Banach-Alaoglu. Existence of a minimum point then follows from Lemma 14.33 and Theorem 14.34. Moreover, if F is strictly convex, suppose that there are two distinct minimizers x and y. Then $F((x + y)/2) < (F(x) + F(y))/2$ which is a contradiction. Hence the minimum is unique in this case. □

Example.

Now we apply this theorem in a concrete situation. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, and let $H_0^1(\Omega)$ be the (real) Hilbert space that we studied in the context of solving the Dirichlet problem

(there we called it S): that is, we take the real C^1 functions on Ω vanishing at the boundary, and complete in the L^2 norm derived from the inner product

$$\langle U, V \rangle = \int_\Omega \nabla U(x) \cdot \nabla V(x)\, dx.$$

This is the Sobolev space of order 1 consisting of functions defined on Ω with one derivative in L^2 and vanishing at the boundary. Recall that we proved the Poincaré inequality

$$\|v\|_2 \leq (\operatorname{diam} \Omega) \|v\|_{H_0^1(\Omega)}$$

for $v \in H_0^1(\Omega)$.

Now consider the nonlinear functional

$$F(v) = \int_\Omega |\nabla v(x)|^2/2 + f(x)v(x)\, dx$$

where $f \in L^2(\Omega)$. This is convex since $t \mapsto t^2$ and $t \mapsto at$ are both convex functions on $\mathbb{R}$. It is continuous on $H_0^1(\Omega)$ since

$$\begin{aligned}|F(v) - F(w)| &\leq \frac{1}{2}\Big|\|v\|^2_{H_0^1(\Omega)} - \|w\|^2_{H_0^1(\Omega)}\Big| + \|f\|_2 \|v - w\|_2 \\ &\leq \frac{1}{2}\|v - w\|_{H_0^1(\Omega)}(\|v\|_{H_0^1(\Omega)} + \|w\|_{H_0^1(\Omega)}) + C\|f\|_2 \|v - w\|_{H_0^1(\Omega)}\end{aligned}$$

Finally, we have $F(v) \to \infty$ as $\|v\|_{H_0^1(\Omega)} \to \infty$, since

$$\begin{aligned}|F(v)| &\geq \frac{1}{2}\|v\|^2_{H_0^1(\Omega)} - C\|f\|_2 \|v\|_{H_0^1(\Omega)} \\ &\geq \frac{1}{2}\|v\|^2_{H_0^1(\Omega)} - \frac{C}{2}\Big(\epsilon\|v\|^2_{H_0^1(\Omega)} + \epsilon^{-1}\|f\|_2^2\Big)\end{aligned}$$

for any $\epsilon > 0$. So we can choose $\epsilon = 1/2C$, say.

The point of considering this function F is that a minimizer u of F is a weak solution of the PDE

$$\Delta u = f \text{ on } \Omega, \quad u = 0 \text{ at } \partial\Omega. \tag{14.5}$$

Indeed, by differentiating F at the minimum point, we find that for any $v \in H_0^1(\Omega)$,

$$\frac{d}{dt}F(u + tv) = 0 \implies \int_\Omega \Big(\nabla u(x) \cdot \nabla v(x) + f(x)v(x)\Big)\, dx = 0.$$

Theorem 14.37 then says that F attains its minimum. Moreover, it is not hard to see that F is strictly convex, so the minimum is unique. Therefore, the equation (14.5) has a unique weak solution for every $f \in L^2(\Omega)$.

Example. The previous example gave the (weak) solution to a linear PDE. However, one great strength of this approach is that we can also solve some nonlinear PDEs with almost equal ease. As an example, we can take instead of the F above,

$$G(v) = \int_\Omega |\nabla v(x)|^2/2 + v(x)^4/4 + v(x)f(x)\, dx.$$

A minimizer of this function is a weak solution to the PDE

$$\Delta u = u^3 + f.$$

The function G is convex, but it may take the value $+\infty$ so cannot be continuous. However, G is norm lower semicontinuous. To see this, suppose that we have a sequence of functions $f_n \in H_0^1(\Omega)$ converging in norm to f. We need to show that $G(f) \leq \liminf G(f_n)$. Define $a = \liminf G(f_n)$. By passing to a subsequence, we can assume that $G(f_n) \to a$.

Then

$$\int_\Omega |\nabla f_n(x)|^2\, dx \to \int_\Omega |\nabla f(x)|^2\, dx \tag{14.6}$$

by convergence in norm. As for the quartic term, we note that $\|f_n - f\|_{H_0^1(\Omega)} \to 0$ implies that $\|f_n - f\|_2 \to 0$. By again passing to a subsequence, we can assume that $f_n \to f$ a.e.. Therefore, $f_n^4 \to f^4$ a.e.. Under this assumption, Fatou's lemma implies that

$$\int f^4 \leq \liminf \int f_n^4.$$

Combining this with (14.6) shows that $G(f) \leq a$, as required.

Theorem 14.37 shows that G has a unique minimizer. Therefore there is a unique weak solution to the equation $\Delta u = u^3 + f$ with $u \in H_0^1(\Omega)$.

Example. Before we do the next example, we need the following classical result.

Theorem 14.38 (Rellich Compactness theorem)**.** *The inclusion map from $H_0^1(\Omega)$ to $L^2(\Omega)$ is compact.*

Now, as a variation of the examples above, consider the problem of minimizing the function F above, with $f \equiv 0$,

$$F(v) = \int_\Omega |\nabla v(x)|^2/2 + f(x)v(x)\, dx,$$

on the subset S of $H_0^1(\Omega)$ consisting of vectors with L^2 norm equal to 1.

Take a minimizing sequence $u_n \in S$, that is such that $\lim F(u_n) = \inf_{v \in S} F(v)$. By passing to a subsequence, we may assume that u_n converge weakly to an element $u \in H_0^1(\Omega)$, and we will have $F(u) \leq \lim F(u_n) = \inf_{v \in S} F(v)$.

There is an apparent flaw in this argument, namely that u might not lie in S, in which case we have not found a minimizer in S. But in fact, I claim that u *does* lie in S.

Lemma 14.39. *If u_n converge weakly in $H_0^1(\Omega)$, and $T : H_0^1(\Omega) \to L^2(\Omega)$ is a compact operator, then Tu_n converge strongly in $L^2(\Omega)$.*

Since T is the inclusion map, Tu_n all have L^2 norm 1, and therefore z must have L^2 norm 1. Therefore, u has L^2 norm 1, so is an element of S, as claimed.

The minimizer u satisfies

$$\frac{d}{dt} F(u + tv)\Big|_{t=0} = 0 \text{ for all } v \in S \text{ such that } u \perp v \text{ in } L^2(\Omega).$$

This means that

$$u \perp v \text{ in } L^2(\Omega) \implies \int_\Omega \nabla u(x) \cdot \nabla v(x)\, dx = 0.$$

Thus, for arbitrary v, $v - \langle v, u \rangle u$ is perpendicular to u, so

$$\int_\Omega \nabla u(x) \cdot \nabla \Big(v(x) - \langle v, u \rangle u \Big)\, dx = 0.$$

Thus, u solves the equation

$$\Delta u = -\lambda u$$

(weakly) where $\lambda = \int |\nabla u|^2 > 0$. This is a way of finding an eigenfunction of Δ corresponding to its smallest eigenvalue.

As a final application, we can apply the same to the functional $G(v) = \int |\nabla v|^2/2 + v^4/4$ restricted to the set of $v \in H_0^1(\Omega)$ where $\|v\|_2 = 1$. We find a minimizer which solves a 'nonlinear eigenvalue problem'

$$\Delta u + u^3 = -\lambda u.$$