Functional Analysis Notes

Hang, Spring 2009

Table of contents

Texts:

1. Lax, *Functional Analysis*

2. Reed-Simon, *Methods of Modern Mathematical Physics, Vol. 1*

3. Yosida, *Functional Analysis*

4. Rudin, *Functional Analysis*

5. Conway, (GTM) *A Course in Functional Analysis*

Some motivational remarks about Functional Analysis? Linear spaces, linear operators in infinite dimensions requires much more machinery. A large focus on linear function spaces and various applications. "Functional" refers to a linear function from $X \to \mathbb{K}$ (X is linear space, $\mathbb{K}$ is some scalar field) and such functionals will play a central role in studying function spaces.

Hahn-Banach Theorems

Given a real vector space $X/\mathbb{R}$ (notation), a subspace $X_0 \subset X$ and a linear functional l_0 defined on X_0, it is easy to extend l_0 to a linear functional l on X (for instance, if we decompose $X = X_0 + Y$ and define $l(x_0 + y) = l_0(x_0)$). Usually l_0 has an additional property that we want to preserve in the extension.

Even though we have restricted to real vector spaces, the results carry through to complex vector spaces with only a slight alteration.

Theorem 1. *Let $X/\mathbb{R}$ be a real vector space and $X_0 \subset X$ a subspace, with l_0 a linear functional defined on X_0. Suppose that $l_0(x) \le p(x)$ for all $x \in X_0$, where $p: X \to \mathbb{R}$ satisfying*

- *Positive Homogeneity.* $p(\lambda x) = \lambda p(x)$ *for* $\lambda > 0$
- *Subadditivity.* $p(x+y) \le p(x) + p(y)$

*Then there exists a **linear** functional l defined on X such that $l|_{X_0} = l_0$ and $l \le p$. (i.e. l is an extension of l_0 that is also dominated by p)*

Remark: p can be relaxed to be just convex as well.

Proof. First show you can extend by one dimension while preserving the domination property, e.g. from X_0 to $X_0 + \operatorname{span}\{y\}$ for $y \notin X_0$. Then turn the set of all possible extensions into a poset and apply Zorn's lemma (existence of maximal linearly ordered chain), and finally show that there is a largest element in the maximal linearly ordered chain. ☐

Example 2. **(Extension of Positive Functionals)** Let E be a set and

$$X = \{f: E \to \mathbb{R} \text{ s.t. } \sup_{x \in E} |f(x)| < \infty\}$$

Assume $Y \subset X$ is a linear subspace that contains an element h such that $h(x) \ge c_0 > 0$ (bounded away from zero), then if $l_0: Y \to \mathbb{R}$ is a linear functional that is positive (i.e. for all $g \in Y$ such that $g \ge 0$, we have that $l_0(g) \ge 0$), then l_0 can be extended to a linear functional l defined on X that is also positive.

The proof involves finding an appropriate p to apply Hahn Banach. Use

$$p(f) := \inf_{\substack{g \ge f \\ g \in Y}} l_0(g)$$

and prove that p is positive homogeneous, subadditive. Then Hahn Banach shows that for any negative $f \in X$ ($f \le 0$), $l(f) \le p(f) \le 0$. For the last inequality note that $0 \in Y$ so that $p(f) \le p(0) = 0$.

Separation of Convex Sets

Theorem 3. *Let $X/\mathbb{R}$ be a real vector space, and $A, B \subset X$ convex with $A \cap B = \varnothing$, A has an **absorbing point**. Then there exists a nonzero linear functional $l: X \to \mathbb{R}$ such that $l(x) \le l(y)$ for all $x \in A$ and $y \in B$.*

We say that x_0 is an **absorbing point** of A if for any x, we can find a scalar λ such that $x \in x_0 + \lambda(A - x_0)$. That is, when A is scaled relative to x_0, for sufficiently large (or small) scalars every point is included (absorbed).

Remark 4. When 0 is an absorbing point of a convex set $A \subset X$, we can define the Minkowski functional (gauge) p_A by

$$p_A(x) = \inf_{\substack{\lambda > 0 \\ x \in \lambda A}} \lambda$$

which is a useful positive homogeneous, subadditive function.

Example 5. If $A = B_1$, then $p_A(x) = |x|$.

The theorem can be reduced to the following lemma:

Lemma 6. *Let $A \subset X$ be a convex subset with an absorbing point, and let $x_0 \notin A$. Then there exists a nonzero linear functional $l: X \to \mathbb{R}$ such that $l|_A \leq l(x_0)$.*

Proof. (of Theorem given Lemma) Use $A - B$, which still has an absorbing point, and since A, B are disjoint, $0 \notin A - B$, and thus for some linear functional l we have that $l|_{A-B} \leq l(0) = 0$, or in other words, $l(x) \leq l(y)$ for $x \in A$ and $y \in B$. □

Proof. (of Lemma) Without loss of generality assume 0 is an absorbing point of A (otherwise shift it). Then we can take the Minkowski functional of A. Then define a linear functional on the subspace spanned by x_0, $\mathbb{R}x_0$, by $l_0(tx_0) = p_A(tx_0)$. This means in particular that $l_0(tx_0) \leq p_A(tx_0)$, and we can extend to a linear functional on X such that $l(x) \leq p_A(x)$. Then for any $x \in A$, we have that

$$l(x) \leq p_A(x) \leq 1 = p_A(x_0) = l(x_0)$$

□

Generalization of Hahn Banach

Theorem 7. *Let $X/\mathbb{R}$ be a real vector space and let X_0 be a subspace. Let $l_0: X_0 \to \mathbb{R}$ be a linear functional on X_0 and $p: X \to \mathbb{R}$ a positive homogeneous and subadditive function with $l_0 \leq p|_{X_0}$ as before. Additionally, let $\mathcal{A}$ be a collection of linear maps from $X \to X$ such that*

1. *For all $A \in \mathcal{A}$, $AX_0 \subset X_0$ (i.e. X_0 is invariant under A)*
2. *For all $A \in \mathcal{A}$, $l_0(Ax) = l_0(x)$ for all $x \in X_0$.*
3. *For all $A, B \in \mathcal{A}$, $AB = BA$ (commutative linear maps)*
4. *For all $A \in \mathcal{A}$, $p(Ax) = p(x)$ for all $x \in X$*

Then there exists an extension to a linear functional $l: X \to \mathbb{R}$ such that $l|_{X_0} = l_0$, $l \leq p$ and $l(Ax) = l(x)$ for $x \in X$.

Example 8. From real analysis (Vitali) we know that there does not exist a nontrivial finite measure on the power set of the unit circle $\mathbb{S}^1$ which is rotation invariant. However, if we relax the condition of countable additivity to just finite additivity, we can use the theorem above to show that there does exist a finitely-additive rotation-invariant set function $\mu: P(\mathbb{S}^1) \to [0, \infty)$.

Let $X = \{u: \mathbb{S}^1 \to \mathbb{R}, \sup_{z \in \mathbb{S}^1} |u(z)| < \infty\}$, $\mathcal{A} = \{R_\theta: \theta \in \mathbb{R}\}$ where $R_\theta u(z) = u(e^{i\theta}z)$. Let Y be the subset of X consisting of the Lebesgue-measurable functions. Then on Y we can define $l_0(u) = \int_{\mathbb{S}^1} u d\theta$, where we note that for Lebesgue measurable sets A, the Lebesgue measure is recovered by $\lambda(A) = l_0(\mathbf{1}_A)$. Also l_0 is rotation-invariant $l_0(R_\theta u) = l_0(u)$ by the rotation-invariance of the Lebesgue measure. Now to extend to arbitrary subsets, we simply extend l_0 to a rotation-invariant functional l defined on all of X and define $\mu(A) = l(\mathbf{1}_A)$, and this gives us a finitely additive (linear) rotation-invariant set function as desired.

The only thing that remains is to find a positive-homogeneous and subadditive function p that is also rotation-invariant and that dominates l. Since $l_0(u) \leq \int_{\mathbb{S}^1} u d\theta \leq 2\pi |u|_\infty$, we can use $p(u) = 2\pi |u|_\infty$, which is subadditive and positive homogeneous and dominates l_0.

Proof. First turn $\mathcal{A}$ into an algebra by using $\mathcal{A}_1 = \{A_1 A_2 \cdots A_m: A_i \in \mathcal{A}\} \supset \mathcal{A}$. Then examine the convex hull $\mathcal{C} = \mathrm{co}(\mathcal{A})$ (set of convex combinations of $\mathcal{A}$), and use the functional

$$q(x) = \inf_{A \in \mathcal{C}} p(Ax)$$

noting that $q(x) \le p(x)$. The reason q is useful is because $q(x - Ax) \le 0$ and $q(Ax - x) \le 0$:

$$q(x - Ax) \le \frac{1}{n}p[(I + A + \ldots + A^{n-1})(I - A)x] = \frac{1}{n}p((I - A^n)x) \le \frac{p(x) + p(-x)}{n} \to 0$$

Note we need q to be positive homogeneous and subadditive. This follows from

$$p\left(\frac{A+B}{2}x + \frac{A+B}{2}y\right) \le \frac{p(Ax) + p(Bx)}{2} + \frac{p(Ay) + p(By)}{2}$$

and taking infimum over $A, B \in \mathcal{C}$ gives the result.

Got A so that $p(Ax)$ is close to inf, $p(By)$ is close to inf.... bound below by $p(Ax + By)$

the same computation works for $q(Ax - x)$ as well. This additional property shows us that the extension to $l \le q$ satisfies $l(x - Ax) \le 0$ and $l(Ax - x) \le 0$ so that $l(x) = l(Ax)$ as desired. □

Hahn-Banach for $\mathbb{C}$

Given a real-valued linear functional $l_1: X \to \mathbb{R}$ there exists a unique complex-valued functional $l: X \to \mathbb{C}$ such that $l_1 = \operatorname{Re}\{l\}$, and also $l(x) = l_1(x) - i l_1(ix)$. Then we can prove the following version of Hahn-Banach for complex scalar field:

Theorem 9. *Let $X/\mathbb{C}$ be a vector space over $\mathbb{C}$, and $X_0 \subset X$ a subspace. Let $l_0: X_0 \to \mathbb{C}$ be a complex valued linear functional over X_0, and $p: X \to \mathbb{R}$ be positive homogeneous ($p(\alpha x) = |\alpha| p(x)$) and subadditive, with $|l_0| \le p$ on X_0. Then there exists an extension to a linear functional $l: X \to \mathbb{C}$ such that $l|_{X_0} = l_0$ and $|l(x)| \le p(x)$.*

Proof. Apply Hahn-Banach to $\operatorname{Re} l_0$ to get an extension to the real functional l_1 over X. Then there is a unique functional l with real part l_1, satisfying the requirements. □

Banach Spaces

Preliminaries: norm, seminorm, Cauchy sequence, complete.

A Banach space is a complete normed linear space. Examples are $C(X, \mathbb{R})$ the space of continuous functions from a compact, Hausdorff space X to $\mathbb{R}$, and $L^p(X, \mu)$ for $1 \le p \le \infty$ (to show below).

Hölder's Inequality

Quick proof: $|fg|_1 \le |f|_p |g|_q$ for $\frac{1}{p} + \frac{1}{q} = 1$, $1 \le p, q \le \infty$. The case $p = 1$, $q = \infty$ is immediate. In general, without loss of generality $|f|_p = |g|_q = 1$ by scaling. By concavity of $\ln x$, we note that

$$|f(x)||g(x)| \le \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

so that integrating gives $|fg|_1 \le 1$ as desired.

This inequality can be generalized slightly as well. First note that $|f^\alpha|_p = |f|_{\alpha p}^\alpha$ by definition. Then suppose that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Then

$$|fg|_{L^r} = ||f|^r |g|^r|_{L^1}^{1/r} \le \big||f|^r\big|_{L^{p/r}}^{1/r} \big||g|^r\big|_{L^{q/r}}^{1/r} = |f|_{L^p} |g|_{L^q}$$

and inductively, we have that if $\frac{1}{r} = \frac{1}{p_1} + \ldots + \frac{1}{p_m}$, then

$$|f_1 \cdots f_m|_{L^r} \leq |f_1|_{L^{p_1}} \cdots |f_m|_{L^{p_m}}$$

Duality Principle

Let p' denote the conjugate exponent satisfying $\frac{1}{p} + \frac{1}{p'} = 1$. Assume $(X,\, \mu)$ is a σ-finite measure space. Then for any $f\colon X \to \mathbb{C}$,

$$|f|_p = \sup_{\substack{|g|_{p'} = 1 \\ fg \,\text{integrable}}} \left| \int_X fg \, d\mu \right|$$

Proof. An upper bound on $\int fg$ is obtained by Hölder, the other constructs an explicit $g \in L^{p'}$ that achieves the upper bound, $g = \varphi \frac{f^{p/p'}}{|f|_p^{p/p'}}$ where $|\varphi| = 1$ and $\varphi |f| = f$. The case where $|f|_p = \infty$ uses σ-finiteness for $X = \bigcup_k X_k$, $Y_n = \bigcup_{k=1}^n X_k$, and $f_n = \min(f, n)\mathbf{1}_{Y_n}$. Then RHS $\geq |f_n|_p \to \infty$ □

Generalized Minkowski

Theorem 10. *Let $1 \leq p \leq \infty$ and $(X, \mu), (Y, \nu)$ be σ-finite. Then if $F\colon X \times Y \to [0, \infty]$, then*

$$\left(\int_X \left(\int_Y F(x,y)\, d\nu(y) \right)^p d\mu(x) \right)^{1/p} \leq \int_Y \left[\int_X F(x,y)^p \, d\mu(x) \right]^{1/p} d\nu(y)$$

Or in other words,

$$\left\| \int_Y F(\,\cdot\,, y)\, d\nu(y) \right\|_{L^p(X)} \leq \int_Y \|F(\,\cdot\,, y)\|_{L^p(X)}\, d\nu(y)$$

Proof. This is just using duality principle and Fubini (this requires σ-finiteness):

$$\begin{aligned} \left\| \int_Y F(\,\cdot\,, y)\, d\nu(y) \right\|_{L^p(X)} &= \sup_{\|g(x)\|_{p'} = 1} \left| \int_X g(x) \overline{\int_Y F(x,y)\, d\nu(y)}\, d\mu(x) \right| \\ &\leq \int_Y \int_X |g(x)||F(x,y)|\, d\nu(y)\, d\mu(x) \\ &\leq \int_Y \left(\sup_{\|g(x)\|_{p'} = 1} \int_X |g(x)||F(x,y)|\, d\mu(x) \right) d\nu(y) \\ &= \int_Y \|F(\,\cdot\,, y)\|_{L^p(X)}\, d\nu(y) \end{aligned}$$

□

As a consequence, we have the following fact

Proposition 11. *Assume $0 < p \leq q < \infty$, and suppose $F \geq 0$. Then*

$$\Big| |F(x,y)|_{L^p(X)} \Big|_{L^q(Y)} \leq \Big| |F(x,y)|_{L^q(Y)} \Big|_{L^p(X)}$$

Proof. First we note that $|f^\alpha|_p = |f|_{\alpha p}^\alpha$ by definition. Then

$$\begin{aligned}
\left||F(x,y)|_{L^p(X)}\right|_{L^q(Y)} &= \left||F^p|_{L^1(X)}^{1/p}\right|_{L^q(Y)} \\
&= \left||F^p|_{L^1(X)}\right|_{L^{q/p}(Y)}^{1/p} \\
&\leq \left||F^p|_{L^{q/p}(Y)}\right|_{L^1(X)}^{1/p} \\
&= \left||F^p|_{L^{q/p}(Y)}^{1/p}\right|_{L^p(X)} \\
&= \left||F|_{L^q(Y)}\right|_{L^p(X)}
\end{aligned}$$

where the inequality follows from Generalized Minkowski and the fact that $q/p \geq 1$. □

From this fact, we can derive the Minkowski inequalities:

Theorem 12. *For $1 \leq p \leq \infty$, we have that*

$$|f+g|_p \leq |f|_p + |g|_p$$

and for $0 < p \leq 1$, we have that if f, g are nonnegative, then

$$|f+g|_p \geq |f|_p + |g|_p$$

Proof. This follows from the previous proposition using an appropriately defined $F(x,y)$. Choose $Y = \{1, 2\}$ with ν counting measure, and $F(x,1) = |f(x)|$ and $F(x,2) = |g(x)|$, and apply the previous proposition with exponents p and 1.

For $1 \leq p \leq \infty$, we have

$$|f+g|_{L^p} \leq ||f(x)| + |g(x)||_{L^p} = \left||F|_{L^1(Y)}\right|_{L^p(X)} \leq \left||F|_{L^p(X)}\right|_{L^1(Y)} = |f|_{L^p} + |g|_{L^p}$$

and for $0 < p \leq 1$, and nonnegative f, g, we have the same sequence of steps except that we need nonnegativity for the very first equality below:

$$|f+g|_{L^p} = ||f(x)| + |g(x)||_{L^p} = \left||F|_{L^1(Y)}\right|_{L^p(X)} \geq \left||F|_{L^p(X)}\right|_{L^1(Y)} = |f|_{L^p} + |g|_{L^p}$$

□

Thus in particular, $|\cdot|_{L^p}$ is not a norm when $p < 1$, and that for $1 \leq p \leq \infty$, $L^p(X, \mu)$ is a normed linear space.

Completeness of L^p

To show that L^p is complete for $1 \leq p \leq \infty$, given a Cauchy sequence f_k, pass to a subsequence so that $|f_k - f_{k+1}|_{L^p} < 2^{-i}$. Then noting $f_m(x) = f_1(x) + \sum_{k=1}^m (f_{k+1}(x) - f_k(x))$ we show that the sum converges a.e. x by showing that the sum is absolutely convergent for a.e. x. Let

$$g(x) = |f_1(x)| + \sum_{k=1}^m |f_{k+1}(x) - f_k(x)|$$

and note that

$$|g(x)|_{L^p} \leq |f_1|_{L^p} + \sum_{k=1}^{m} |f_{k+1} - f_k|_{L^p} \leq |f_1|_{L^p} + 1$$

so that $g \in L^p$. This implies that $g(x) < \infty$ for a.e. x so that the sum is absolutely convergent for a.e. x. Thus $\lim_{m \to \infty} f_m(x)$ exists a.e. x, and denote the limit by $f(x)$. We show that $f_m \to f$ in L^p:

$$\lim_{m \to \infty} \int |f_m - f|^p = \int \lim_{m \to \infty} |f_m - f| = 0$$

by dominated convergence, since $|f_m - f|^p \leq (|f_m| + |f|)^p \leq (2|g|)^p \leq 2^p |g|^p$ which is integrable.

Examples of Banach Spaces

Example 13. $C(X, \mathbb{R})$ with the supremum norm, where X is compact, Hausdorff. Here completeness follows from the fact that uniform limit of continuous functions is continuous and the fact that if f_n is Cauchy in $C(X, \mathbb{R})$, then $f_n(x)$ is Cauchy in $\mathbb{R}$.

Example 14. $L^p(X, \mu)$ where (X, μ) is a σ-finite measure space, and $1 \leq p \leq \infty$.

Example 15. Let $\Omega \subset \mathbb{R}^n$ be open, bounded.

- $C^m(\Omega) = \{u \colon \Omega \to \mathbb{R},\, u \text{ has continuous partial derivatives up to order } m\}$
- $C^m(\overline{\Omega}) = \{u \in C^m(\Omega),\ \text{the partial derivatives of } u \text{ have continuous extension to } \overline{\Omega}\}$

Denoting $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\partial^\alpha u = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$, $|\alpha| = \alpha_1 + \ldots + \alpha_n$ (multi-index notation), we have seminorms

$$p_k(u) = \sum_{|\alpha| = k} |\partial^\alpha u|_\infty$$

and from these we can construct a norm (not unique) by

$$|u|_{C^m(\overline{\Omega})} = \sum_{k=0}^{m} p_k(u)$$

These two spaces equipped with the norm above are Banach spaces.

Example 16. Let (X, d) be a metric space. Then for $0 < \alpha \leq 1$, let

$$C^\alpha(X) = \left\{ u \colon X \to \mathbb{R},\ \sup_{x_1 \neq x_2} \frac{|u(x_1) - u(x_2)|}{d(x_1, x_2)^\alpha} < \infty \right\}$$

be the space of Hölder-continuous functions of order α. Then we have a seminorm

$$[u]_\alpha = \sup_{x_1 \neq x_2} \frac{|u(x_1) - u(x_2)|}{d(x_1, x_2)^\alpha}$$

which we use to construct the norm

$$|u|_{C^\alpha(X)} = |u|_\infty + [u]_\alpha$$

This is a Banach Space.

Example 17. Let $\Omega \subset \mathbb{C}$, and consider for $1 \le p \le \infty$,

$$A^p(\Omega) = \{ f : \Omega \to \mathbb{C}, \text{ holomorphic}, |f|_{L^p} < \infty \} \subset L^p(\Omega)$$

Then $A^p(\Omega)$ equipped with the L^p norm is a Banach space.

Since $A^p(\Omega)$ is a subset of $L^p(\Omega)$, to show completeness it suffices to show that the L^p limit of a sequence of functions in A^p is also in A^p, e.g. that A^p is closed under the L^p metric.

Note that for $p = \infty$, completeness follows from the result that if f_n is a sequence of analytic functions and $f_n \to f$ uniformly on all compact subsets of Ω, then f is analytic.

For $p < \infty$, first note that by the Cauchy Integral Formula, and for r chosen so that $B_r(z_0) \subset \Omega$,

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{it})\, dt$$

Multiplying by r and integrating in r, we have that on $B_R(z_0) \subset \Omega$,

$$\frac{1}{2} R^2 f(z_0) = \frac{1}{2\pi} \int_{B_R(z_0)} f(z)\, dA$$

or

$$f(z_0) = \frac{1}{|B_R(z_0)|} \int_{B_R(z_0)} f(z)\, dA$$

Now note that

$$|f_n(z_0) - f_m(z_0)| \le \frac{1}{|B_R(z_0)|} \int_{B_R(z_0)} |f_n(z) - f_m(z)|\, dA$$

and using Hölder's inequality, we have that

$$|f_n(z_0) - f_m(z_0)| \le \frac{1}{|B_R(z_0)|^{1-1/p'}} \|f_n - f_m\|_p$$

If we consider compact subsets $K \subset \Omega$, then there is a uniform constant for which

$$|f_n(z_0) - f_m(z_0)| \le C \|f_n - f\|_p,\ z_0 \in K$$

(by compactness). Therefore, $\|f_n - f_m\|_{L^\infty(K)} \to 0$ on compact K, so that we have upgraded the convergece so that $f_n \to f$ uniformly on all compact subsets of K, which means that f is analytic (apply theorem that says that $\{f_n\}$ is a normal family of functions, etc...). A detail missing here is that focusing on K, f_n converges uniformly to some g, but then we show that $f_n \to g$ in L^p as well, using

$$\int_K |f_n - g|^p \le \|f_n - g\|_{L^\infty(K)}^p \mu(K) \to 0$$

and by uniqueness of limits $g = f$ on K. This shows that $f_n \to f$ uniformly on compact sets.

Example 18. Hardy space $1 \le p \le \infty$. Define

$$H^p(B_1) = \left\{ f\colon B_1 \to \mathbb{C} \text{ holomorphic such that } \sup_{0<r<1} \int_0^{2\pi} |f(re^{i\theta})|^p \, d\theta < \infty \right\}$$

with the norm

$$|f|_{H^p} = \left(\sup_{0<r<1} \int_0^{2\pi} |f(re^{i\theta})|^p \, d\theta \right)^{1/p}$$

Then H^p with the norm is a Banach space.

Example 19. Let $\Omega \subset \mathbb{R}^n$ and fix $m \in \mathbb{N}$. Let $1 \le p \le \infty$ and define

$$E_m = \left\{ u \in C^m(\Omega) \text{ such that } |\partial^\alpha u|_{L^p(\Omega)} < \infty \text{ for all } |\alpha| \le m \right\}$$

Then define

$$|u|_{E_m} = \left(\sum_{|\alpha| \le m} |\partial^\alpha u|_{L^p(\Omega)}^p \right)^{1/p}$$

Then $(E_m, |\cdot|_{E_m})$ is not complete. Its completion will be denoted $W^{m,p}(\Omega)$ (Sobolev Space).

For instance, in the case $m = 0$ it is easy to find a sequence of continuous functions that converge in the L^p norm to a discontinuous function.

Completion of a Normed Linear Space

Given a normed linear space X, a completion of X means a pair (φ, Y) where Y is complete, $\varphi\colon X \to Y$ is a linear norm-preserving (and thus injective) map, and $\overline{\varphi(X)} = Y$.

Example 20. If $X = C([0,1])$ with the L^1 norm, then we know that $(C([0,1]), L^1)$ is not complete. The completion is $(f \mapsto f, L^1)$, identifying X as elements of L^1, which is complete, and noting that the subspace of continuous functions are dense in L^1.

Proposition 21. *Every normed linear space has a unique completion (Though for practical purposes, it is crucial to have a concrete description)*

Proof. The reason is that we can take Z to be the set of all Cauchy sequences in X, which forms a vector space. Then define a seminorm $|\{x_k\}_{k=1}^\infty|_Z = \lim_{k\to\infty} |x_k|$, which is subadditive and positive homogenous. Quotient out the subspace for which $|\{x_k\}|_Z = 0$, i.e. $Z_0 = \{(x_k) \in Z\colon \lim x_k = 0\}$, then Z/Z_0 is a Banach space, and the map $\varphi\colon X \to Z/Z_0$ given by $\varphi(x) = (x, x, x, \ldots) + Z_0$ gives the completion.

To show uniqueness, suppose (φ_1, Y_1) and (φ_2, Y_2) are two completions of X. Then we show that there exists a norm-preserving isomorphism between Y_1 and Y_2. Note that the images $\varphi_1(X)$ and $\varphi_2(X)$ are dense in Y_1, Y_2 respectively, and φ_1 is invertible as a map from $X \to \varphi_1(X)$, a linear isomorphism. Then we define $\Phi\colon \varphi_1(X) \to \varphi_2(X)$ by $\Phi(\varphi_1 x) = \varphi_2 x$, noting that $|\varphi_1 x| = |x| = |\varphi_2 x|$, so Φ is a linear, norm-preserving isomorphism. We then extend by continuity. Given $y \in Y_1$, we choose any sequence $\varphi_1 x_n \in \varphi_1(X)$ converging to y, and define

$$\Phi(y) = \lim_{n\to\infty} \Phi(\varphi_1 x_n)$$

where the limit exists because $|\Phi(\varphi_1 x_n) - \Phi(\varphi_1 x_m)| = |\varphi_1 x_n - \varphi_1 x_m| \to 0$ so that $\Phi(\varphi_1 x_n)$ is Cauchy. This definition is well defined since if $\varphi_1 x_n' \to y$ then $\lim_{n\to\infty} \Phi(\varphi_1 x_n - \varphi_1 x_n') = 0$ (the norm goes to 0 with the same computation). Note $|\Phi(y)| = \lim_{n\to\infty} |\Phi(\varphi_1 x_n)| = \lim_{n\to\infty} |\varphi_1 x_n| = |y|$ so that Φ is a linear norm preserving map from $Y_1 \to Y_2$. The same construction can be used to construct a linear norm preserving map Ψ from $Y_2 \to Y_1$. Then $\Psi\Phi y = y$ essentially by construction... □

Separability

A linear metric space (X, d) is called **separable** if there exists a dense, countable subset, i.e. exists a set E with countably many points such that $\overline{E} = X$. There is the following fact:

Proposition 22. *(X, d) is separable if and only if X has a countable base, i.e. there exists a countable collection of open sets $\mathcal{B}$ such that for all open sets U, U can be expressed as a union of sets in $\mathcal{B}$.*

Proof. If (X, d) is separable, then take balls with rational radii centered around each point in the dense subset. Conversely, given a countable base, take an arbitrary point in each set of the base and denote the collection by E. Then, for any open ball, there exists an element of the base contained inside, which contains a point of E, and thus E is dense. □

Properties:

- If (X, d) separable, and $Y \subset X$, then (Y, d) is separable.
- If (X, d) separable, and $\bigcup_{\alpha\in\Lambda} U_\alpha = X$ is an open cover, then there exists $\Lambda_0 \subset \Lambda$, a countable set of subindices such that $\bigcup_{\alpha\in\Lambda_0} U_\alpha = X$.

  To see this, let $\mathcal{B}$ be a countable base. Every U_α is a union (at most countable) of elements in $\mathcal{B}$. Let $\mathcal{B}'$ consist of all the sets $B \in \mathcal{B}$ for which $B \subset U_\alpha$ for some $\alpha \in \Lambda$. Note that the union of all elements in $\mathcal{B}'$ is also X. Then for each $B \in \mathcal{B}'$, choose any $\alpha_B \in \Lambda$ for which $B \subset U_{\alpha_B}$. Then the desired subcover is $\{U_{\alpha_B}, B \in \mathcal{B}'\}$.

Examples of Separable Spaces

- $(C[0,1], |\cdot|_\infty)$ is separable (polynomials with rational coefficients)
- $(L^p(\Omega), |\cdot|_{L^p})$ is separable for $\Omega \subset \mathbb{R}^n$ (step functions, continuous functions, etc)
- $(L^\infty(\Omega), |\cdot|_{L^\infty})$ is not separable for Ω with infinitely many elements. (biject to the sequence of 0's and 1's)

A Variational Problem

Let X be a Banach space, and K be a convex, closed subset of X, and let $y \notin K$. We wish to find $x_0 \in K$ that minimizes $|x - y|$ for $x \in K$, i.e. $|x_0 - y| \leq |x - y|$ for all $x \in K$. Does such a point exist? And if so, how many points?

In $\mathbb{R}^n$, the answer is simple. Let $d = \inf_{x\in K} |x - y|$. By definition of inf, we can find a sequence $x_n \in K$ such that $|x_n - y| \to d$. Noting that $|x_n| \leq |x_n - y| + |y| \leq d + |y|$ so that the sequence is bounded, and by sequential compactness of bounded sequences in $\mathbb{R}^n$, there exists a subsequence $x_{n'}$ converging to some $x_0 \in K$. Then $|x_0 - y| = d$. Here we used the assumption that K is closed and the Heine-Borel property, that every bounded sequence has a convergent subsequence. Unfortunately this does not work in infinite dimensions.

Proposition 23. *If X is a normed linear space, and every bounded sequence has a convergent subsequence, then* $\dim X < \infty$.

The proof follows quite directly from the following lemma:

Lemma 24. *If Y is a proper, closed subspace of X, then for any $\varepsilon > 0$, there exists $x \in X$ such that $|x| = 1$ and $\inf_{y \in Y} |x - y| \geq 1 - \varepsilon$.*

Given this lemma, and the fact that every finite dimensional space is closed, we can construct a sequence so that $|x_n| = 1$ and $\inf_{y \in \operatorname{span}\{x_1, \ldots, x_n\}} |x_{n+1} - y| \geq \frac{1}{2}$, so that $d(x_j, x_k) \geq \frac{1}{2}$. Note at the induction step we apply the lemma to $Y = \operatorname{span}\{x_1, \ldots, x_n\}$ and $\varepsilon = \frac{1}{2}$. Such a sequence cannot have a convergent subsequence, but is bounded.

Proof. (of Lemma) First we note that X / Y can be equipped with the norm

$$\|x + Y\| = \inf_{y \in Y} |x - y|$$

for $x + Y \in X / Y$. For any $y_1, y_2 \in Y$, we have

$$\inf_{y_1 + y_2 \in Y} |x_1 + x_2 - (y_1 + y_2)| \leq |x_1 - y_1| + |x_2 - y_2|$$

taking inf over y_1 and y_2 gives subadditivity. Suppose $\inf_{y \in Y} |x - y| = 0$. Then there exists a sequence y_n such that $|x - y_n| \to 0$, and since Y is closed, we have that $x \in Y$ so that $x + Y = 0 + Y$ (the zero vector in X/Y). (Thus, if Y is not closed, $\|\cdot\|$ is just a seminorm).

Also, if X is complete, so is X/Y. This is because if $x_n + Y$ is Cauchy, then $\|x_n - x_m + Y\| = \inf_{y \in Y} |x_n - x_m - y| \to 0$, and can find $y_m \in Y$ such that $|(x_n - y_n) - (x_m - y_m)| \to 0$, so that $x_n - y_n$ is Cauchy, and by completeness $x_n - y_n \to z$, and so $x_n + Y \to z + Y$.

Now there exists $x_0 + Y \in X / Y$ such that $\|x_0 + Y\| = 1$. For $\varepsilon > 0$, there exists y such that

$$1 \leq |x_0 + y| < 1 + \varepsilon$$

Then let $x = \frac{x_0 + y}{|x_0 + y|}$, so that $|x| = 1$, and furthermore,

$$\inf_{y \in Y} |x - y| = \|x + Y\| = \frac{\|x\|}{|x_0 + y|} \geq \frac{1}{1 + \varepsilon}$$

as desired. ∎

Remark 25. As an aside, the fact that all finite dimensional subspaces are closed is a consequence of the completeness of $L^p(\{1, \ldots, n\}, \text{counting})$ and the fact that all norms are equivalent in finite dimensional vector spaces. Let $V = \mathbb{R}^n$ and $|\cdot|_V$ be a norm for V, and let $|\cdot|_2$ be the standard Euclidean norm.

1. First we show that $|\cdot|_V$ is continuous on $(\mathbb{R}^n, |\cdot|_2)$. Note

$$|x|_V = \left| \sum_{k=1}^n x_k e_k \right|_V \leq \sum_{k=1}^n |x_k| |e_k|_V \leq \left(\sum_{k=1}^n |e_k|_V \right) |x|_2 = C |x|_2$$

Then $\big| |x|_V - |y|_V \big| \leq |x - y|_V \leq C |x - y|_2$ so that $|\cdot|_V$ is continuous on $(\mathbb{R}^n, |\cdot|_2)$.

2. Then, let $b = \min_{|x|_2=1} |x|_V > 0$, noting that $\{x\colon |x|_2 = 1\}$ is a compact set, and thus $|\cdot|_V$ maps this set to a compact set in $\mathbb{R}$. Since 0 is not in the image ($|x|_V = 0$ only for $x = 0$, which has 2-norm 0), it must be the case that $b > 0$. Then for any nonzero x, we have that

$$\left|\frac{x}{|x|_2}\right|_V \geq b$$

or that $|x|_V \geq b|x|_2$

These two imply that $b\,|x|_2 \leq |x|_V \leq C\,|x|_2$ for two constants b, C. This implies that the topologies induced by $|\cdot|_V$ and $|\cdot|_2$ are the same.

Strict and Uniform Convexity

Back to the problem about minimizing $|x - y|$ over $x \in K$ where $y \notin K$ is fixed. K is convex and closed subset of X. First we show that the problem does not always have a solution.

Example 26. (Minimum not achieved) Let $X = C[0,1]$ and $|f|_\infty = \max_{x \in [0,1]} |f(x)|$ the ∞ norm, and let

$$K = \left\{ f \in C[0,1],\ \int_0^1 f(x)\,dx = 1, f(0) = 0, f(1) = 1 \right\}$$

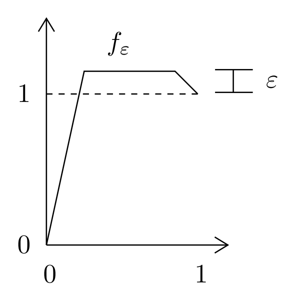
This is a convex, closed (uniform convergence on compact set allows swapping limits) set. Note that $0 \notin K$, and $d = \inf_{f \in K} |f - 0|_\infty$ is not achieved.

First, we show that $d = 1$.

For any $f \in K$, we note that

$$1 = \int_0^1 f(x)\,dx \leq |f|_\infty$$

so that $d \geq 1$. To show $d \leq 1$ we consider specific $f \in K$. Consider the function f_ε constructed by modifying the constant function $f(x) = 1$ by forcing $f_\varepsilon(0) = 0$ and compensating for the loss in the integral by raising the maximum value of the function slightly to $1 + \varepsilon$, i.e.



This can be chosen so that $\int f_\varepsilon = 1$. Then we have that $d \leq \|f_\varepsilon\|_\infty = 1 + \varepsilon$ and since ε is arbitrary, we have that $d \leq 1$.

Now we show that d is not achieved. Suppose that $f \in K$ and $\|f\|_\infty = 1$, so that $\int_0^1 f(x)\,dx = 1$. Then $\int_0^1 (1 - f(x))\,dx = 0$, which means that $1 - f(x) \equiv 0$ so $f(x) = 1$, contradicting $f \in K$.

The issue here is compactness. We can take a sequence whose norms tend to 1, but this is a bounded sequence which may not have a convergent subsequence to some element in K (if it did, then the limit has norm 1)

Next we show that the solution may not even be unique.

Example 27. (Solution not unique) In $\mathbb{R}^2$, with $|x|_\infty = \max(|x_1|, |x_2|)$ let K be the unit ball (which is the square $[-1,1]^2$) and $y = (2,0)$. Then $y \notin K$, the minimum distance $d = \inf_{x\in K} |x-y| = 1$, but every point $(1,t)$, $|t| \leq 1$ achieves this distance.

The condition for the uniqueness of the solution is **strict convexity**.

X is **strictly convex** if for all $x, y \in X$,

$$|x| = 1, |y| = 1, \left|\frac{x+y}{2}\right| = 1 \Longrightarrow x = y$$

Proposition 28. *If X is strictly convex, and $x, y \in X$, $x \neq 0, y \neq 0$, and $|x+y| = |x| + |y|$, then $x = \lambda y$ for some $\lambda > 0$.*

Proof. Note we may write the above as

$$\left|\frac{|x|}{|x|+|y|}\frac{x}{|x|} + \frac{|y|}{|x|+|y|}\frac{y}{|y|}\right| = 1$$

Define $f(t) = \left|t\frac{x}{|x|} + (1-t)\frac{y}{|y|}\right|$ for $t \in [0,1]$. Note that by the strict convexity of X, if $\frac{x}{|x|} \neq \frac{y}{|y|}$, then f is also strictly convex so that $f(t) < 1$ for $t \in (0,1)$ $(f(0) = f(1) = 1)$. This is because the definition of strict convexity tells us that two *different* points on a line cannot have the same norm since if the midpoint of x, y has the same norm as x and y.

Given the claim, then the above shows that $f\left(\frac{|x|}{|x|+|y|}\right) = 1$ whereas $\frac{|x|}{|x|+|y|} \in (0,1)$. This implies that $\frac{x}{|x|} = \frac{y}{|y|}$ and $x = \frac{|x|}{|y|}y$ as desired. □

Proposition 29. (Uniqueness) *Let X be a Banach space, $K \subset X$ be strictly convex and closed, and $y \notin K$. Suppose $d = dist(y,K) > 0$, and $x_0 \neq x_1$ satisfies $|x_0 - y| = |x_1 - y| = d$. Then $x_0 = x_1$.*

Proof. Without loss of generality, we may assume $y = 0$ by translation invariance. Then we note that $|x_0| = |x_1| = d$. This implies that x_0 is not a multiple of x_1, so that by Proposition 28,

$$\left|\frac{x_0 + x_1}{2}\right| < \frac{|x_0| + |x_1|}{2} = d = \operatorname{dist}(0,K)$$

But this is a contradiction since by convexity $\frac{x_0+x_1}{2} \in K$ but its distance from $y = 0$ is smaller than d. Thus we conclude that $x_0 = x_1$. □

For existence, we need to talk about a still stronger notion of convexity. X is called **uniformly convex** if for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $|x| = |y| = 1$,

$$\left|\frac{x+y}{2}\right| > 1 - \delta \Longrightarrow |x - y| < \epsilon$$

(i.e. if the midpoint of two unit vectors is close to a unit vector, then the two unit vectors are close)

Remark 30. Let $(X, |\cdot|)$ be strictly convex. If $\dim X < \infty$. then it is also uniformly convex.

Proof. Suppose X is finite dimensional but not uniformly convex. Then there exists $\epsilon > 0$ such that for all $\delta > 0$, we can find unit vectors x_n, y_n such that $\left|\frac{x_n + y_n}{2}\right| > 1 - \delta$ but $|x_n - y_n| > \epsilon$. Since X is finite dimensional and we have a bounded sequence, we can find a subsequence $x_{n'}$, $y_{n'}$ where $x_{n'}$, $y_{n'}$, and $\frac{x_{n'} + y_{n'}}{2}$ all converge to x_∞, y_∞, $\frac{x_\infty + y_\infty}{2}$. But this implies that $\left|\frac{x_\infty + y_\infty}{2}\right| = 1$, and $|x_\infty - y_\infty| \geq \epsilon$, which means that X is not strictly convex. ☐

Proposition 31. *Let $(X, |\cdot|)$ be uniformly convex. Then it is strictly convex.*

Proof. Let $|x| = |y| = 1$ such that $\left|\frac{x+y}{2}\right| = 1$ Let $\epsilon > 0$, and δ be chosen as in the definition of uniform convexity. Then since $\left|\frac{x+y}{2}\right| = 1 > 1 - \delta$, we have that $|x - y| < \epsilon$. Since ϵ is arbitrary, we have that $|x - y| = 0$ and thus $x = y$. ☐

This tell us that if a solution exists to the variational problem in the case that X is uniformly convex, then the solution is unique. Now we can show existence.

Theorem 32. *Let X be a uniformly convex Banach space, and let $K \subset X$ be convex and closed. Then for all $y \in X$, there exists a unique $x_0 \in K$ such that*

$$|y - x_0| = \inf_{x \in K} |y - x| = dist(y, K)$$

Proof. First we note that if $y \in K$, the theorem is trivial. Let $y \notin K$. Then let

$$b = \inf_{x \in K} |y - x| > 0$$

By the definition of b, we can find a sequence $x_n \in K$ such that $|x_n - b| \rightarrow b$. We now show that this sequence is Cauchy, in which case we have found the vector of minimal distance to y.

Now we make the following observation: By uniform convexity, if we have two sequences x_n, y_n such that $|x_n| \rightarrow b$, $|y_n| \rightarrow b$, and $\left|\frac{x_n + y_n}{2}\right| \rightarrow b$, then $|x_n - y_n| \rightarrow 0$. This follows from looking at the unit vectors $\frac{x_n}{|x_n|}$, $\frac{y_n}{|y_n|}$ and

$$\left|\frac{\frac{x_n}{|x_n|} + \frac{y_n}{|y_n|}}{2}\right| \rightarrow 1$$

and applying the definition of uniform convexity.

Note that

$$b \leq \left|\frac{x_n + x_m}{2} - y\right| \leq \frac{|x_n - y|}{2} + \frac{|x_m - y|}{2} \rightarrow b \text{ as } n, m \rightarrow \infty$$

Thus

$$\left|\frac{x_n - y}{2} + \frac{x_m - y}{2}\right| \rightarrow b, \ |x_n - y| \rightarrow b, \ |x_m - y| \rightarrow b$$

so that by our observation

$$|(x_n - y) - (x_m - y)| \rightarrow 0$$

Therefore the sequence x_n is Cauchy, converging to x_∞ where

$$|x_n - y| \rightarrow |x_\infty - y| = b = \text{dist}(y, K)$$

The uniqueness of the solution follows from Proposition 29 and the fact that uniform convexity implies strict convexity. □

Remark 33. For a simpler proof using weak convergence with the assumption that X is reflexive, see Example 71, in the section about lower semicontinuity.

Uniformly Convex Spaces

What sort of spaces are uniformly convex?

Theorem 34. *$L^p(X)$ is uniformly convex for $1 < p < \infty$*

It is easy to see (in $\langle R \rangle^2$ say) that in the cases $p = 1$ and $p = \infty$ the space is not even strictly convex (the unit circles in the respective norms are composed of straight lines). The theorem follows easily from Clarkson inequality (not proved).

Theorem 35. (Clarkson Inequality) *Let $u, v \in L^p(X)$. Then we have the following inequalities:*

1. *If $1 \leq p < 2$, then*

$$\left|\frac{u+v}{2}\right|_{L^p}^q + \left|\frac{u-v}{2}\right|_{L^p}^q \leq \left(\frac{|u|_{L^p}^p + |v|_{L^p}^p}{2}\right)^{q-1}$$

2. *If $2 \leq p < \infty$, then*

$$\left|\frac{u+v}{2}\right|_{L^p}^p + \left|\frac{u-v}{2}\right|_{L^p}^p \leq \left(\frac{|u|_{L^p}^q + |v|_{L^p}^q}{2}\right)^{p-1}$$

In particular, for $p = 2$, this becomes

$$|u+v|_{L^2}^2 + |u-v|_{L^2}^2 \leq 2(|u|_{L^2}^2 + |v|_{L^2}^2)$$

Proof. For reference, see Adams, *Sobolev Spaces*. □

The proof of uniform convexity from the Clarkson inequality is now easy. If $|u|, |v| = 1$, then the RHS is 1, and thus if $\left|\frac{u+v}{2}\right|$ is near 1, then by the inequality $\left|\frac{u-v}{2}\right|$ is near 0.

Another example of uniformly convex spaces are Hilbert spaces.

Hilbert Spaces

A Hilbert space is a complete inner product space. There is the Cauchy Schwarz inequality

$$|\langle u, v \rangle| \leq |u||v|$$

and the norm is $|x| = \sqrt{\langle x, x \rangle}$.

Examples are L^2 with the inner product $\langle f, g \rangle = \int fg$.

Proposition 36. *Every Hilbert space is uniformly convex.*

Proof. From properties of inner product, we have the parallelogram equality:

$$|x+y|^2+|x-y|^2=2(|x|^2+|y|^2)$$

If we now write this as

$$\left|\frac{x+y}{2}\right|^2+\left|\frac{x-y}{2}\right|^2=\frac{|x|^2+|y|^2}{2}$$

then we see that if $|x|=|y|=1$, the RHS is 1, and thus if $\left|\frac{x+y}{2}\right|^2\to 1$ then $\left|\frac{x-y}{2}\right|^2\to 0$ and $|x-y|\to 0$. □

Orthogonal Decomposition

Proposition 37. *Let H be a Hilbert space over the real or complex scalar field, and let $X\subset H$ be a closed subspace. Then $H=X\oplus X^{\perp}$ where*

$$X^{\perp}=\{y\in H\colon\langle x,y\rangle=0\}$$

Furthermore, $X^{\perp}$ is a closed subspace of H, by the continuity of the inner product.

Proof. First note that $X\cap X^{\perp}=\{0\}$ since if $x\in X\cap X^{\perp}$ then $\langle x,x\rangle=0$ so that $x=0$.

Then, we show that $H=X+X^{\perp}$. Let $z\in H$. Since every Hilbert Space is uniformly convex, using Theorem 32, there exists a $x_0\in X$ such that $|x_0-z|\le|x-z|$ for all $x\in X$. We show that $x\perp z-x_0$ for all $x\in X$ so that $z=x_0+(z-x_0)$ where $x_0\in X$ and $z-x_0\in X^{\perp}$.

If we consider

$$\varphi(t)=|z-(x_0+tx)|^2=|z-x_0|^2-2t\,\mathrm{Re}\,\langle z-x_0,x\rangle+t^2|x|^2$$

then the minimum is achieved when $t=0$, and so

$$0=\varphi'(0)=-2\mathrm{Re}\langle z-x_0,x\rangle$$

Thus $\mathrm{Re}\langle z-x_0,x\rangle=0$ for all $x\in X$. This implies that $\langle z-x_0,x\rangle=0$ if we consider both $x,ix\in X$. □

Properties: Let H be a Hilbert space over real or complex scalar field.

- Let S be any subset of H. $S^{\perp}=(\overline{\mathrm{span}(S)})^{\perp}$ and $S^{\perp}$ is closed.

  Proof. Note that by linearity, if $\langle x,z\rangle=\langle y,z\rangle=0$, then $\langle\alpha x+\beta y,z\rangle=0$, and thus if $z\in S^{\perp}$ then $z\in(\mathrm{span}(S))^{\perp}$, so $S^{\perp}\subset(\mathrm{span}(S))^{\perp}$. Since opposite inclusion is immediate, $S^{\perp}=(\mathrm{span}(S))^{\perp}$.

  If $y_n\in\mathrm{span}(S)$ converges to some y in $\overline{\mathrm{span}(S)}$, and $\langle y_n,z\rangle=0$, then by the continuity of the inner product $\langle y,z\rangle=0$ also. This implies the result. □

- Let $X\subset H$ be a closed subspace. Then $(X^{\perp})^{\perp}=X$.

  Proof. First, if $x\in X$, then for all $y\in X^{\perp}$, $\langle x,y\rangle=0$. Thus $x\in(X^{\perp})^{\perp}$ and so $X\subset(X^{\perp})^{\perp}$

Next, we use the decomposition $H = X \oplus X^{\perp}$. If $z \in (X^{\perp})^{\perp}$ then write $z = x + y$ where $x \in X$ and $y \in X^{\perp}$. Then

$$0 = \langle z, y \rangle = \langle x, y \rangle + |y|^2 = |y|^2$$

so that $y = 0$ and $z \in X$. Thus $(X^{\perp})^{\perp} \subset X$ and we have the result. □

- Let $X \subset H$ be any subspace. Then $(X^{\perp})^{\perp} = \overline{X}$.

  Proof. Use $(X^{\perp})^{\perp} = (\overline{X}^{\perp})^{\perp} = \overline{X}$. □

Riesz Representation Theorem

Denote the dual space of a vector space X by $X' = \mathcal{L}(H, \mathbb{K})$ where $\mathbb{K}$ is the scalar field. Note that for any $y \in H$ we can define a linear functional in the dual space by $l_y(x) = \langle x, y \rangle$. This is linear, and by Cauchy-Schwarz, $|l_y(x)| \leq |x||y|$ so that the operator norm is $|l_y| = |y|$. It turns out that all linear functionals arise in this manner.

Theorem 38. *Let H be a real or complex Hilbert space, and let l be a linear functional in H'. Then there exists a unique vector $y \in H$ such that $l(x) = \langle x, y \rangle$.*

Proof. Sketch: Consider $N_l^{\perp}$, which is one dimensional since $N_l^{\perp} \cong H / N_l \cong \mathbb{K}$. Then take the unit vector in $N_l^{\perp}$. □

Application: This can be used to provide a proof of the Radon-Nikodym theorem and the Lebesgue Decomposition.

Theorem 39. *Let μ, ν be two finite measures on X. Then ν can be decomposed as*

$$\nu = \nu_{\text{ac}} + \nu_{\text{sing}}$$

where $\nu_{\text{ac}} \ll \mu$ and $\nu_{\text{sing}} \perp \mu$, i.e. there exists a nonnegative integrable function g such that

$$\nu_{\text{ac}}(A) = \int_A g \, d\mu$$

and a set $N \subset X$ such that

$$\mu(N) = \nu_{\text{sing}}(X \backslash N) = 0$$

Also, the decomposition is unique.

Proof. Let $H = L^2(X, \mu + \nu)$. Then define $l(f) = \int_X f d\nu$, and note that

$$|l(f)| \leq \int_X |f| d\nu \leq \int_X d(\mu + \nu) \leq |f|_{L^2(X, \mu + \nu)} \sqrt{(\mu + \nu)(X)}$$

so that $l \in H'$. By the Riesz Representation theorem, there exists $h \in H$ such that

$$\int_X f d\nu = \int_X f h \, d(\mu + \nu)$$

The claim is that $0 \le h \le 1$ for a.e. $\mu+\nu$. To see this, let $E = \{h \ge 1\}$. Then note that using $f = \mathbf{1}_E$,

$$\int_E d(\mu+\nu) \ge \int_E d\nu = \int_E h\, d(\mu+\nu)$$

and so

$$\int_E (h-1)\, d(\mu+\nu) \le 0$$

Since $h > 1$ on E, we see that $(\mu+\nu)(E) = 0$. Likewise, for $F = \{h < 0\}$, we have

$$0 \le \int_F d\nu = \int_F h\, d(\mu+\nu)$$

and since $h < 0$ on F we have that $(\mu+\nu)(F) = 0$. Now assume without loss of generality that $0 \le h \le 1$ everywhere (modifying h on a set of measure zero does not affect the integral).

The intuition here (formal computation) is that "$d\nu = h\, d\mu + h\, d\nu$ so that $d\nu = \frac{h}{1-h} d\mu$", which is okay if $h < 1$. Thus, let us take the set $N = \{h = 1\}$.

Note that if $f = \mathbf{1}_N$, $\nu(N) = \mu(N) + \nu(N)$ so that $\mu(N) = 0$. Then we set

$$\begin{aligned} \nu_{\text{sing}}(A) &= \nu(A \cap N) \\ \nu_{\text{ac}}(A) &= \nu(A \backslash N) \end{aligned}$$

Now we show that $\nu_{\text{ac}}(A) = \int_A g\, d\mu$ where

$$g(x) = \begin{cases} \frac{h(x)}{1-h(x)} & x \in X \backslash N \\ 0 & x \in N \end{cases}$$

(since $\mu(N) = 0$ the value of g on N does not matter, and $h < 1$ for $x \in X \backslash N$). From the definition of ν_{sing} and the usual approximation steps (simple to nonnegative measurable functions) we note that

$$\int f\, d\nu_{\text{ac}} = \int_{N^c} f\, d\nu$$

so that for any nonnegative meausrable $f \ge 0$,

$$\begin{aligned} \int f\, d\nu_{\text{ac}}(A) &= \int_{A \backslash N} f\, d\nu \\ &= \int_{A \backslash N} f h\, d(\mu+\nu) \\ &= \int_A f h\, d\mu + \int_{A \backslash N} f h\, d\nu_{\text{ac}} + \int_{A \backslash N} f h\, d\nu_{\text{sing}} \\ &= \int_A f h\, d\mu + \int_A f h\, d\nu_{\text{ac}} \end{aligned}$$

noting that $\nu_{\text{ac}}(N) = 0$ and $\nu_{\text{sing}}(A \backslash N) = \nu(N \cap A \backslash N) = \nu(\varnothing) = 0$. Then we have that as in the intuition,

$$\int_A f(1-h)\, d\nu_{\text{ac}} = \int_A f h\, d\mu$$

Now take $f = \frac{1}{1-h}\mathbf{1}_{N^c}$, which is nonnegative and measurable. Then

$$\nu_{\text{ac}}(A) = \int_A d\nu_{\text{ac}} = \int_A g d\mu$$

as desired, and in particular g is μ-integrable.

For uniqueness, suppose we have two decompositions

$$\nu = \nu_{\text{ac}} + \nu_{\text{sing}} = \nu'_{\text{ac}} + \nu'_{\text{sing}}$$

where ν_{sing} is supported on $X \setminus N$ and ν'_{sing} on $X \setminus N'$, ν_{ac}, $\nu'_{\text{ac}} \ll \mu$. If we then consider $N_0 = N \cup N'$, we have that

$$\nu_{\text{sing}}(A) = \nu_{\text{sing}}(A \cap N_0) = \nu(A \cap N_0) - \nu_{\text{ac}}(A \cap N_0) = \nu(A \cap N_0)$$

noting that $\mu(N_0) \leq \mu(N) + \mu(N') = 0$ so that $\nu_{\text{ac}}(A \cap N_0) = \int_{A \cap N_0} g d\mu \leq \int_{N_0} g d\mu = 0$. The same holds for $\nu'_{\text{sing}}, \nu'_{\text{ac}}$ so that $\nu_{\text{sing}} = \nu'_{\text{sing}}$ and $\nu_{\text{ac}} = \nu'_{\text{ac}}$. □

Fun List of Applications:

- Radon-Nikodym and Lebesgue Decomposition
- Dirichlet Forms

Adjoints

Let $A \in \mathcal{L}(H_1, H_2)$ where H_1, H_2 are Hilbert spaces.

Fixing $y \in H_2$, $x \mapsto \langle Ax, y \rangle_{H_2}$ is a bounded linear functional on H_1 since $|\langle Ax, y \rangle| \leq |A| |x|_{H_1} |y| H_2$. Thus by Riesz representation there exists some element $A^* y \in H_2$ such that

$$\langle Ax, y \rangle = \langle x, A^* y \rangle$$

for all $x \in H_1$. This allows us to define a bounded linear operator A^* from $H_2 \to H_1$ which maps $y \mapsto A^* y$. (Linearity is straightforward to prove)

Properties:

- $|A^*| = |A|$

  Proof:

  $$|A^*| = \sup_{|y| \leq 1} |A^* y| = \sup_{\substack{|x| \leq 1 \\ |y| \leq 1}} |\langle x, A^* y \rangle| = \sup_{\substack{|x| \leq 1 \\ |y| \leq 1}} |\langle Ax, y \rangle| = \sup_{|x| \leq 1} |Ax| = |A|$$

- $A^{**} = A$

  Proof:

  $$\langle Ax, y \rangle = \langle x, A^* y \rangle = \langle A^{**} x, y \rangle$$

  for all x, y, and thus $A = A^{**}$.

- $(AB)^* = B^* A^*$

Proof:

$$\langle x, ABy\rangle = \langle A^*x, By\rangle = \langle B^*A^*x, y\rangle$$

for all x, y, thus by definition of adjoint, $(AB)^* = B^*A^*$.

- $R_A^\perp = N_{A^*}$, $\overline{R_A} = N_{A^*}^\perp$

  Proof:

$$\begin{aligned} y \in R_A^\perp &\iff \langle y, Ax\rangle = 0 \text{ for all } x \in H_1 \\ &\iff \langle A^*y, x\rangle = 0 \text{ for all } x \in H_1 \\ &\iff A^*y = 0 \\ &\iff y \in N_{A^*} \end{aligned}$$

  Thus $R_A^\perp = N_{A^*}$. Taking the $\perp$ of both sides gives the second statement.

Example 40. (Integral Operators) Let $K \in L^2(\mu \times \nu)$ where (X, μ) and (Y, ν) are measure spaces, and define $A: L^2(\mu) \to L^2(\nu)$ by

$$(Af)(y) = \int_X K(x, y)\, f(x)\, d\mu(x)$$

K is called the **kernel** of the integral operator.

Then A is a bounded linear operator, $|A| \leq |K|_{L^2(\mu\times\nu)}$ and $A^*: L^2(\nu) \to L^2(\mu)$ is also an integral operator with kernel $K^*(y, x) = \overline{K(x, y)}$.

(easy computation)

Orthonormal Bases

Let H be a Hilbert space. Then $(e_\alpha)_{\alpha\in A}$ is called an **orthonormal system** if $\langle e_\alpha, e_\beta\rangle = \delta_{\alpha\beta}$ for $\alpha, \beta \in \Lambda$. It is an **orthonormal basis** if in addition $\{e_\alpha, \alpha \in \Lambda\}^\perp = 0$, i.e. $\overline{\operatorname{span}\{e_\alpha\}} = H$.

Examples:

- $L^2(\mathbb{S}^1)$, the exponentials $e_k = e^{2\pi i k\theta}$ is an orthonormal basis.

- Let $B_1 = \{z \in \mathbb{C}: |z| < 1\}$, $A^2(B_1) = \{f \text{ holomorphic on } B_1, \|f\|_2 < \infty\}$ with inner product

$$\langle f, g\rangle = \int_{B_1} f\,\bar{g}\, d\mu$$

  Then $c_k z^k$ is an orthonormal base (c_k an appropriate normalization constant)

Proposition 41. *Every Hilbert space has at least one orthonormal basis.*

Proof. Use Hausdorff Maximality Principle (Zorn's lemma) on the poset of orthonormal systems with the inclusion ordering ($\{e_k\} \leq \{f_k\} \iff \{e_k\} \subset \{f_k\}$). Clearly every linearly ordered set has an upper bound (take the union), and the maximal element B is a basis since otherwise $B^\perp \neq 0$ so we can add another element to B preserving orthonormality, and this contradicts the maximality of B. $\square$

Proposition 42. *If H is separable, then every orthonormal set is countable.*

Proof. Given an orthonormal system $\{e_\alpha\}$, we note that for $\alpha \neq \beta$, $|e_\alpha - e_\beta| = \sqrt{2}$, and so taking a sufficiently small ball around each e_α gives mutually disjoint balls $\{B_\alpha, \alpha \in \Lambda\}$. Then given a countable dense subset $\{x_n\}$, each B_α necessarily contains at least some point in $\{x_n\}$. Thus we have a correspondence between B_α and a subset of $\{x_n\}$, and since $\{x_n\}$ is countable, there are countably many balls, and thus $\{e_\alpha\}$ is countable. $\square$

Proposition 43. *Let H be a separable Hilbert space, and $\{e_k\}$ and orthonormal basis. Then*

1. *For all $x \in H$, $x = \sum_k \langle x, e_k \rangle e_k$*

2. *(Bessel) $|x|^2 = \sum_k |\langle x, e_k \rangle|^2$*

Proof. To prove (1), consider $x - y = \sum_{k=1}^n \langle x, e_k \rangle e_k$ for some y, then taking the norm squared and taking limits gives

$$\sum_k |\langle x, e_k \rangle|^2 \leq |x|^2$$

(Bessel's inequality). This shows that $\sum_k \langle x, e_k \rangle e_k$ is absolutely summable, and since H is separable, this means that it is summable, i.e. converges to some limit z. Note

$$\langle x - z, e_j \rangle = \langle x, e_j \rangle - \sum_k \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0$$

Since this is true for all e_j, $x - z = 0$ so $x = z$. For (2), apply the norm squared to partial sums in (1), and take the limit. $\square$

Dual Spaces

Given a normed linear space X, the dual space $X' = \mathcal{L}(X, \mathbb{K})$ is the space of linear functionals, and under the operator norm $|l| = \sup_{|x| \leq 1} |l(x)|$ the dual space $(X', |\cdot|)$ is a Banach space.

Proof. Let l_n be a Cauchy sequence in $(X', |\cdot|)$. $|l_n - l_m| \to 0$. Then for a given $x \in X$, $|l_n(x) - l_m(x)| \leq |l_n - l_m||x| \to 0$, so that $l_n(x)$ is Cauchy. This allows us to define a linear functional l by limits,

$$l(x) = \lim_{n \to \infty} l_n(x)$$

which is linear by the linearity of limits. Then, $\frac{|l_n(x) - l(x)|}{|x|} \to 0$ for all $x \neq 0$, so that $|l_n - l| \to 0$, so that l_n converges to l. $\square$

Note that X is not required to be complete.

Recall Definitions: Let X be a locally compact, Hausdorff space.

- $\tau = \{U \text{ open}\}, \mathcal{B} =$ smallest σ-algebra containing τ.

- **Borel measure** is a measure μ on $\mathcal{B}$ such that $\mu(K) < \infty$ for K compact.

- **Regular Borel measure** is a Borel measure such that

$$\mu(E) = \inf_{\substack{U \text{ open} \\ E \subset U}} \mu(U) = \sup_{\substack{K \text{ compact} \\ K \subset E}} \mu(K)$$

  i.e. measure can be approximated from above by open sets and below by compact sets.

Examples:

1. In Hilbert Spaces, $H' \cong H$ by Riesz Representation.

2. For $1 \leq p < \infty$ and (X, μ) a σ-finite measure space, $L^p(\mu)' \cong L^{p'}(\mu)$, where $\frac{1}{p} + \frac{1}{p'} = 1$. The isomorphism is given by $\theta: L^{p'} \to (L^p)'$ defined by

$$\theta g(f) = \int f g \, d\mu$$

   This mapping is norm-preserving by Hölder's inequality and duality principle, and is onto by the Radon-Nikodym theorem (need to show that the resulting function is in $L^{p'}$.

3. Let X be compact and Hausdorff, then $C(X, \mathbb{C})$, the space of continuous functions from $X \to \mathbb{C}$ is a Banach space under the supremum norm, and

$$C(X, \mathbb{C})' \cong \{\mu: \mu \, \text{regular complex Borel measure on} \, X\}$$

   with norm $\|\mu\| = |\mu|(X)$, where $|\mu|$ is the total variation measure of μ. Here the isomorphism is given by

$$\theta\mu(\varphi) = \int_X \varphi \, d\mu$$

4. (Generalizes (3)). Let X be locally compact (every point has a compact neighborhood) and T_2. Then

$$C_0(X, \mathbb{C}) = \{\varphi: X \to \mathbb{C} \, \text{bounded continuous function}, \forall \varepsilon > 0, \exists K \subset X \, \text{compact}, |\varphi| \leq \varepsilon \, \text{on} \, X \backslash K\}$$

   is a Banach space under the supremum norm, and as above,

$$C_0(X, \mathbb{C})' \cong \{\mu: \mu \, \text{regular complex Borel measure on} \, X\}$$

Remark 44. We will show (4) using the subspace C_c, which is the space of continuous functions with compact support. This is a normed linear space under the supremum norm, but it is not a Banach space. The dual space is more complicated here...

Also, if X is locally compact metric space, then every Borel measure is regular. Regularity is important because this implies that $C_c(X) \subset L^1(X)$ is dense.

We will make use of the following theorem:

Theorem 45. *Let X be a locally compact Hausdorff space, and let $l: C_c(X, \mathbb{R}) \to \mathbb{R}$ be a positive linear functional, i.e. $l(\varphi) \geq 0$ for $\varphi \geq 0$. Then there exists a unique regular Borel measure μ such that*

$$l(\varphi) = \int_X \varphi \, d\mu$$

Also, if $l: C_c(X, \mathbb{C}) \to \mathbb{C}$ *is a linear functional such that for all* $\varphi \in C_c(X, \mathbb{R})$, $\varphi \geq 0$,

$$\sup_{\substack{\psi \in C_c(X, \mathbb{C}) \\ |\psi| \leq \varphi}} |l(\psi)| < \infty$$

then there exists a unique regular Borel measure μ *and a measurable function* $\theta: X \to \mathbb{C}$ *with* $|\theta| = 1$ *such that*

$$l(\varphi) = \int_X \varphi \theta d\mu$$

Using this theorem, we see that if $l \in C_0(X)'$, then l is also a linear functional on C_c satisfying the boundedness property of the theorem above since given $\varphi \geq 0$, $|\psi| \leq \varphi$ with $\varphi, \psi \in C_c$, we have that

$$|l(\psi)| \leq |l||\psi|_\infty \leq |l||\varphi|_\infty$$

and thus there exists θ, μ with $|\theta| = 1$ such that

$$l(\varphi) = \int_X \varphi \theta d\mu$$

and it remains to justify that μ is finite so that $\theta d\mu$ is a complex measure. Since X is locally compact, Hausdorff, μ is regular, and so taking a seqeuence of compact sets K_n that increase to X, we have that

$$\mu(X) = \lim_n \int \mathbf{1}_{K_n} d\mu \leq \sup_{\substack{\psi \in C_c(X, \mathbb{C}) \\ |\psi| \leq \varphi}} |l(\psi)| < \infty$$

Existence of nontrivial elements in X'

Proposition 46. *Given a normed linear space* X *with* $Y \subset X$*. If* $l_0 \in Y'$*, then there exists an extension to* X' *which preserves norm, i.e.* $l \in X'$ *such that* $l|_Y = l_0$ *and* $|l|_{X'} = |l_0|_{Y'}$*.*

Proof. Since $|l_0(y)| \leq |l_0||y|$ for $y \in Y$, let $p(x) = |l_0||x|$, subadditive positive homogeneous function. Then since $|l_0| \leq p$, we apply Hahn Banach to obtain an extension to $l \in X'$ with $|l| \leq p$ and $l|_Y = l_0$. Then $|l(x)| \leq |l_0||x|$ shows that $|l| \leq |l_0|$ and since $l = l_0$ on Y, $|l_0| \leq |l|$. ☐

Corollary 47. *X normed linear space,* $Y \subset X$ *with* $\dim Y < \infty$*. Then there exists* Z *closed such that* $X = Y \oplus Z$*. (Useful for Fredholm Theory)*

Proof. Let $(e_1, \ldots, e_n)$ be a basis for Y. Then we can find linear functionals $l_1, \ldots, l_n \in X'$ such that $l_i(e_j) = \delta_{ij}$, noting that $l_i(e_j) = \delta_{ij}$ defines linear functionals on Y' (defined for all of Y by linearity, and is continuous since all norms on finite dimensional spaces are equivalent). Then we can extend to X' using the Proposition above. Let $Z = \bigcap_{j=1}^n N_{l_j}$, the intersection of the null spaces. Then we can express any $x \in X$ as

$$x = \underbrace{\sum_j l_j(x) e_j}_{\in Y} + \underbrace{\left(x - \sum_j l_j(x) e_j\right)}_{\in Z}$$

so that $X = Y + Z$. Also, if $x \in Y \cap Z$, then $x = \sum_j c_j e_j$, but $l_k(x) = c_k = 0$ so that $x = 0$.

Z is closed since the null space of a continuous linear functional is closed, and Z is the intersection of null spaces, and hence is also closed. (null space is inverse image of $\{0\}$, which is closed) $\square$

Duality Principle

Proposition 48. (Duality Principle) *For $y \in X$,*

$$|y| = \sup_{\substack{|l| \le 1 \\ l \in X'}} |l(y)|$$

and furthermore there is an $l \in X'$, $|l| \le 1$ that achieves the supremum, i.e. $|l(y)| = |y|$

Proof. Note that $|l(y)| \le |y|$ when $|l| \le 1$ so that $|y| \ge \text{RHS}$. To get the other inequality, we note we can define a linear functional on $\operatorname{span}\{y\} = \mathbb{C}y$ by $l_0(\lambda y) = \lambda |y|$. Then $|l_0| \le 1$, and we can extend to $l \in X'$ such that $|l| = |l_0| \le 1$, and $|l(y)| = |y|$. Thus $|y| = |l(y)| \le \sup_{|l| \le 1} |l(y)|$. $\square$

Corollary 49. *Let X be a normed linear space, $Y \subset X$ a closed subspace. Then if $x_0 \notin Y$, then there exists $l \in X'$ such that $l(x_0) \neq 0$. More specifically, we can find l so that $|l| \le 1$ and $|l(x_0)| = d(x_0, Y)$.*

Proof. Use X/Y, and $\pi: X \to X/Y$ be the injection $\pi(x) = x + Y \in X/Y$. Then $x_0 + Y \neq 0$, and by the proposition we have an $l' \in (X/Y)'$ such that $|l'| \le 1$, $|l'(x_0 + Y)| = |x_0 + Y|$, and we can define $l \in X'$ by

$$l(x_0) = l'(x_0 + Y)$$

where $|l(x_0)| = |l'(x_0 + Y)| = |x_0 + Y| = d(x_0, Y)$. Furthermore, since $l = l' \circ \pi$, we have that

$$|l(x)| = |l'(\pi(x))| \le |l'||\pi||x|$$

so that $|l| \le 1$. $\square$

Application of Duality to Density Problems

Let X be a normed linear space, and let $S \subset X$ be a subset. We would like to know whether $\operatorname{span}\{S\}$ is dense in X. Then we have the following fact:

Proposition 50. *Suppose that $\{l \in X', l|_S = 0\} = \{0\}$, e.g. that the only linear functionals on X that vanish on S is the zero functional, then $\overline{\operatorname{span} S} = X$.*

Proof. Suppose note, then $Y = \overline{\operatorname{span} S} \neq X$. Pick $x_0 \in X \backslash Y$ (in X but not Y), then using Corollary 49 there exists a linear functional $l \in X'$ such that $l(x_0) \neq 0$ but $l|_S = 0$, which is a contradiction. $\square$

Application: Runge Approximation Theorem

Recall that given $f: B_1 \to \mathbb{C}$ holomorphic, there exists a sequence of polynomials converging uniformly to f on $\overline{B_{1/2}}$ (Taylor).

On the annulus $A_{1,2} = 1 < |z| < 2$, given $f: A_{1,2} \to \mathbb{C}$, there may not exist a sequence of polynomials converging uniformly to f on $\overline{A_{1+\varepsilon, 2-\varepsilon}}$, but there does exist a sequence of rational functions converging uniformly to f on $\overline{A_{1+\varepsilon, 2-\varepsilon}}$ (Laurent).

There is a further generalization to $\operatorname{nbd}(K)$ for K compact.

Theorem 51. (Runge) *Let K compact, $f: K \to \mathbb{C}$ holomorphic, and if $E \subset \mathbb{C} \cup \{\infty\}$ such that every component of $\mathbb{C} \cup \{\infty\} \backslash K$ contains at least one element of E, then there exists rational functions R_j with poles in E such that $R_j \to f$ uniformly on K.*

From this, Taylor is with $K = \overline{B_{1/2}}$, $E = \{\infty\}$ and Laurent is with $K = \overline{A_{1+\varepsilon, 2-\varepsilon}}$ and $E = \{0, \infty\}$.

Proof. We need to show that for all $l \in C(\mathrm{nbd}(K))'$, $l(R) = 0$, where R is the space of rational functions with poles in E, that $l = 0$. Let φ be a holomorphic function from $\mathrm{nbd}(K) \to \mathbb{C}$.

Using a contour $\Gamma \subset \mathrm{nbd}(K) \backslash K$ which has winding number 0 on $\mathbb{C} \backslash \mathrm{nbd}(K)$ and 1 on K, we have that by Cauchy's integral formula,

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\xi)}{\xi - z} d\xi, \, z \in K$$

and by linearity (and dominated convergence),

$$l(f(z)) = \frac{1}{2\pi i} \int_\Gamma f(\xi) l\left(\frac{1}{\xi - z}\right) d\xi$$

Thus it suffices to show that $\varphi(\xi) = l\left(\frac{1}{\xi - z}\Big|_{z \in K}\right) = 0$ for $\xi \in \Gamma$. Note φ is holomorphic on $\mathbb{C} \backslash K$, and for $\xi \in E$, $\xi \neq \infty$ we note that

$$\varphi^{(n)}(\xi) = l\left((-1)^n \frac{n!}{(\xi - z)^{n+1}}\right) = 0, \, \xi \in E$$

But this implies that $\varphi \equiv 0$ on the bounded components of $\mathbb{C} \backslash K$ containing ξ, and in particular for $\xi \in \Gamma$, as desired.

Note if $\infty \notin E$, we are done. Otherwise, we need to consider the unbounded component of $\mathbb{C} \backslash K$ and show that $\varphi = 0$ on this component. Note that $\frac{1}{\xi - z} = \frac{1}{\xi} \cdot \frac{1}{1 - z/\xi} = \frac{1}{\xi} \sum_j \left(\frac{z}{\xi}\right)^j$ for $z \in K$ and $|\xi|$ sufficiently large.

Since $(z/\xi)^j$ has a pole at ∞, $l\left(\frac{1}{\xi - z}\right) = 0$ for ξ sufficiently large. Note that we can modify the contour Γ so that the part in the unbounded component of $\mathbb{C} \backslash K$ is just $|\xi| = R$ for R large enough. $\square$

Dual Variational Problem

Let X be a normed linear space over $\mathbb{R}$ and A a subset. Define

$$S_A(l) = \sup_{x \in A} l(x)$$

Properties:

- If $A \subset B$ then $S_A \leq S_B$

  Proof: This property is immediate: $\sup_{x \in A} l(x) \leq \sup_{x \in B} l(x)$ since the supremum on the RHS is taken over a larger set.

- $S_{A+B} = S_A + S_B$.

  Proof. This is also immediate:

$$\sup_{x \in A, y \in B} l(x + y) = \sup_{x \in A} l(x) + \sup_{y \in B} l(y)$$

since $l(x + y) = l(x) + l(y)$, and taking the supremum over A first, then the supremum over B second gives the equality. $\square$

- $S_A = S_{\mathrm{co}(A)} = S_{\overline{\mathrm{co}(A)}}$

  Proof: The first equality holds since $l(\lambda a + (1 - \lambda)b) = \lambda l(a) + (1 - \lambda)b \leq \max\,(l(a), l(b))$. The second holds by continuity of l.

- If A, B are closed, convex, then $S_A \geq S_B \Longleftrightarrow A \supset B$

  Proof: The first property shows ($\Longleftarrow$) direction. Suppose the other direction does not hold. Then there exists $x_0 \in B \backslash A$ and $r > 0$ such that $B_r(x_0) \cap A = \varnothing$. Then there exists $l \in X'$ such that $l \neq 0$ and $l(x) \leq l(y)$ for $x \in A$, $y \in B_r(x_0)$ (Hahn Banach separation theorem). Now choose z so $l(z) = 1$, and then for ε sufficiently small,

$$l(x) \leq l(x_0 - \varepsilon z) = l(x_0) - \varepsilon$$

  Taking the supremum over x shows that $S_A(l) \leq l(x_0) - \varepsilon < S_B(l)$, which is a contradiction (note, the strict inequality is because this works for any ε sufficiently small).

- Let C be any set. Then $x_0 \in \overline{\mathrm{co}(C)} \Longleftrightarrow l(x_0) \leq S_C(l)$

  Proof:

$$x_0 \in \overline{\mathrm{co}(C)} \Longleftrightarrow l(x_0) = S_{\{x_0\}}(l) \leq S_{\overline{\mathrm{co}(C)}}(l) = S_C(l)$$

Theorem 52. *Let K be a closed, convex set, $x_0 \neq K$. Then*

$$d(x_0, K) = \inf_{y \in K} |x_0 - y| = \sup_{\substack{l \in X' \\ |l| \leq 1}} (l(x_0) - S_K(l))$$

Proof. Note that for $|l| \leq 1$, applying the third property above shows that

$$l(x_0) - S_K(l) \leq l(x_0) - l(y) \leq |x_0 - y|$$

and taking the infimum over $y \in K$ and the supremum over $l \in X'$, $|l| \leq 1$ shows that

$$\inf_{y \in K} |x_0 - y| \geq \sup_{\substack{l \in X' \\ |l| \leq 1}} (l(x_0) - S_K(l))$$

To get the opposite inequality, let $d = d(x_0, K)$, and let $r < d$ be arbitrary. Consider an r-neighborhood of K, i.e. $K + B_r$ ($B_r = B_r(0)$). By the hyperplane separation theorem (Theorem 3) on $K + B_r$ and x_0, we can find a functional l where $l(x) \leq l(x_0)$ for $x \in K + B_r$, and without loss of generality we may choose $|l| \leq 1$ by scaling. This means that

$$S_K(l) + S_{B_r}(l) = S_{K + B_r}(l) \leq l(x_0)$$

Then because $S_{B_r}(l) = \sup_{|x| \leq r} l(x) = r\,|l| = r$, and rearranging, we have that

$$r = S_{B_r}(l) \leq l(x_0) - S_K(l) \leq \sup_{|l| \leq 1} (l(x_0) - S_K(l))$$

and since r is arbitrarily close to $d(x_0, K)$, we have the result.

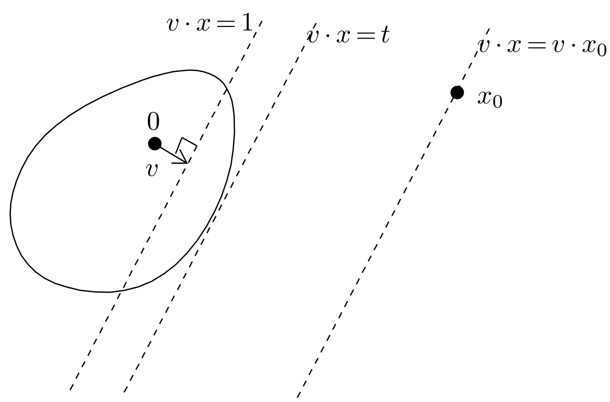
This is in Lax, *Functional Analysis*, Theorem 8.18 $\square$

Remark 53. if K is a closed **subspace**, then from Corollary 49 we can find $l \in X'$ with $|l| \leq 1$, where $l(y) = 0$ for $y \in K$ and $l(x_0) = d(x_0, K)$, and thus $S_K(l) = 0$ and

$$\inf_{y \in K} |x_0 - y| = d(x_0, K) = l(x_0) - S_K(l) \leq \sup_{\substack{l \in X' \\ |l| \leq 1}} (l(x_0) - S_K(l))$$

But this completely ignores the role of S_K.

Remark 54. This is the picture in $\mathbb{R}^2$. The term sup $(l(x_0) - S_K(l))$ probes "directions". Given a direction vector v of unit length, we study the lines formed by $v \cdot x = t$. We increase t until we hit the last point of K. Then we subtract this from $v \cdot x_0$. This gives us a rough estimate (lower bound) of how far x_0 is from K. We then choose the direction v that maximizes this quantity. See picture.



In the picture above, $l(x_0) - S_K(l)$ is $v \cdot x_0 - t$. If we rotate v counterclockwise, we will get the correct distance. Using a nonconvex K in $\mathbb{R}^2$ we can find an example where the theorem is false.

Weak Convergence

Recall that for infinite dimensional normed linear spaces, closed and bounded sets are not necessarily sequentially compact (Heine-Borel property not satisfied) in the topology induced by the norm. However, we still hope for some type of convergence using a different topology.

Definition 55. *X normed linear space, $x_j \in X$. We say $x_j \Longrightarrow x$ weakly if*

$$l(x_j) \rightarrow l(x) \text{ for all } l \in X'$$

(Sometimes will write "$x_j \rightarrow x$ weakly" also...)

We will discuss the topology associated with this notion of convergence later.

Remark 56. If $x_j \rightarrow x$ weakly and $x_j \rightarrow y$ weakly, then $x = y$ (i.e. the topology of weak convergence is Hausdorff). The reason is because then $l(x - y) = 0$ for all l, and by duality principle (Proposition 48) $x - y = 0$.

Example 57. For $1 \leq p < \infty$, $e^{ikx} \to 0$ weakly in $L^p(0, 1)$ by the Riemann-Lebesgue lemma, which says that

$$\int f(x) e^{ikx} \to 0$$

for $f \in L^q$, $1 \leq q \leq \infty$. Note $(L^p)' \cong L^q$ for $1 \leq p < \infty$. The dual of L^∞ is not L^1.

Remark 58. Note that if $x_n \to x$ in norm, i.e. $|x_n - x| \to 0$, then $x_n \to x$ weakly. This follows from

$$|l(x_n) - l(x)| \leq |l| \, |x_n - x| \to 0$$

Even with this weaker notion of convergence, it not always true that bounded sequences have convergent subsequences. The condition for this is that the space be **reflexive**.

Reflexive Spaces

First, we note a natural map from $X \to X''$; that is, given $x \in X$, there is a natural mapping κ: $X \to X''$ which maps x to a linear functional κ_x: $X' \to \mathbb{K}$ given by $\kappa_x(l) = l(x)$. Note that $\|\kappa_x\| = \sup_{|l| \leq 1} |l(x)| = |x|$, so that κ is an isometric embedding of $X \to X''$.

A Banach space X is said to be **reflexive** if $\kappa(X) = X''$ (i.e. that κ is onto).

Facts About Reflexive Spaces:

1. For $1 < p < \infty$, $L^p(X, \mu)$ is reflexive.

   This follows since $(L^p)'' \cong (L^{p'})' \cong L^p$.

2. $L^1(\Omega), \Omega \subset \mathbb{R}^n$ is not reflexive, and neither is $C[0, 1]$.

   These use the following fact:

   Proposition 59. *If X' is separable, then so is X (in their respective norm topologies)*

   Proof. Let $\{l_i, i = 1, 2, ...\}$ be a countable dense subset of X' with $l_i \neq 0$. Then for all i, there exists $x_i \in X$ such that $|x_i| = 1$ and $|l(x_i)| \geq \frac{1}{2}|l_i|$. The claim is that $\overline{\operatorname{span}\{x_i\}} = X$, which we show using Proposition 50. If not, then there exists $l \in X'$ nonzero such that $l(x_i) = 0$ for all i. By density of l_i, given $\varepsilon > 0$ there exists l_i such that $|l_i - l| \leq \varepsilon$. Then

   $$\frac{1}{2}|l_i| \leq |l_i(x_i)| = |l_i(x_i) - l(x_i)| \leq \varepsilon$$

   and thus

   $$|l| \leq |l - l_i| + |l_i| \leq \varepsilon + 2\varepsilon = 3\varepsilon$$

   so that $l = 0$, a contradiction. ☐

   Now given the Proposition, the fact that $L^1(\Omega)$ is not reflexive follows easily, since the dual of $L^1(\Omega)$ is isometrically ismomorphic to $L^\infty(\Omega)$, which is not separable (Ω infinite). Now if L^1 were reflexive, then the double dual would also be L^1, and since L^1 is separable, the previous proposition implies L^∞ is separable, a contradiction.

   The fact that $C[0, 1]$ is not reflexive follows from the fact that $C[0, 1] \subset L^1$, so that the dual $(C[0, 1])' \supset (L^1)' = L^\infty$, and since L^∞ is not separable, neither is the dual of $C[0, 1]$, and thus $C[0, 1]$ is not reflexive, by the same reasoning as for L^1.

3. Closed subspaces of reflexive spaces are reflexive.

   Let X be reflexive, and $Y \subset X$ is a closed subspace. We only need to show that κ: $y \mapsto \kappa_y$ maps onto Y''. Let $\varphi \in Y''$. Then given $l \in X'$, we consider $\varphi(l|_Y) = l(x)$ for some x. Note that for any l such that $l\big|_Y = 0$, we have that $l(x) = \varphi\big(l|_Y\big) = 0$. Since Y is closed, this implies that $x \in Y$ (apply Proposition 50 to $Y \subset Y + \mathbb{K}x$)

   Now for any $l_0 \in Y'$, we can extend arbitrarily to some $l \in X'$ so that $l|_Y = l_0$. Then

   $$\varphi(l_0) = \varphi(l|_Y) = l(x) = l_0(x)$$

   where the last equality follows since $x \in Y$. Thus $\kappa(x) = \varphi$ for $x \in Y$, and so κ is onto.

4. Hilbert spaces are reflexive.

   In Hilbert spaces, by Riesz representation we have that $H \cong H' \cong H''$ (isometric isomorphism follows from duality principle (Proposition 48))

5. Uniformly convex spaces are reflexive. This is proved in the section about Double Polar Theorem (Corollary 117).

Theorem 60. *Let X be a Banach space. Then X is reflexive if and only if every bounded sequence has a weakly convergent subsequence.*

Proof. First suppose that X is reflexive. Let x_i be a bounded sequence in X. Let $Y = \overline{\operatorname{span}\{x_i\}}$. This is a separable subspace of X, and since Y is a closed subspace, Y is also reflexive. Then $Y'' = \kappa(Y)$ is separable and thus Y' is separable. This means that there exists $\{l_j, j = 1, 2, \ldots\}$ which is dense in Y'. Now for each j, $\{l_j(x_i), i = 1, 2, \ldots\}$ is a bounded sequence and therefore has a convergent subsequence. By using diagonal procedure, we can find a single convergent subsequence $x_{i'}$ for which $l_j(x_{i'})$ converges for all j. Define $\varphi: Y' \to \mathbb{K}$ by $\varphi(l_j) = \lim_{i'} l_j(x_{i'})$, and we can extend to Y' by continuity ($\{l_j\}$ is dense in Y').

Then $\varphi \in Y''$, and thus for some $x \in X$, $\kappa(x) = \varphi$, i.e. $\varphi(l) = l(x)$ for all $l \in Y'$. Then since $l(x_{i'}) \to \varphi(l) = l(x)$ for all l, $x_{i'} \to x$ weakly.

The converse is true, but relies on more machinery. □

Application to Galerkin's Method

Galerkin's method is about finding solutions to PDEs by approximation in finite dimensional subspaces.

Let $A_i(x)$ be an $n \times n$ matrix for each $x \in \mathbb{R}^n$ and $1 \leq i \leq n$, and $B(x)$ be an $n \times n$ matrix for each x. For $u: \mathbb{R}^n \to \mathbb{R}$, define L as the operator

$$Lu = \sum_{i=1}^{n} A_i \,\partial_i u + Bu$$

Let A_i, B be smooth and using the standard basis in $\mathbb{R}^n$, $e_1, \ldots, e_n$, suppose $A_i(x + e_j) = A_i(x)$ and $B(x + e_j) = B(x)$ (i.e. periodic).

Given $f: \mathbb{R}^n \to \mathbb{R}$ periodic, we wish to solve $Lu = f$ for some periodic u.

Define

$$H = \{u \in L^2_{\text{loc}}(\mathbb{R}^n): u(x + e_j) = u(x) \forall j\}$$

which is a Hilbert space with the inner product $\langle u, v\rangle = \int_Q uv\,dx$ with Q the unit cube in $\mathbb{R}^n$.

Proposition 61. *If $A_i^T = A_i$ and there exists $\lambda > 0$ such that $B + B^T - \sum_i \partial_i A_i \geq \lambda I$, then for every $f \in H$, there exists $u \in H$ such that*

$$\langle u, L^* v\rangle = \langle f, v\rangle$$

for every $v \in C^1 \cap H$. (this condition says that u is a weak solution to $Lu = f$).

Notational Warning: For convenience we will start writing $A_i \partial_i$ in place of $\sum_i A_i \partial_i$. It is taken to be understood that we are summing over i.

Proof. A simple computation using the definition $\langle u, L^* v\rangle = \langle Lu, v\rangle$ and integration by parts shows that

$$L^* v = (-A_i \partial_i - \partial_i A_i + B^T) v$$

Since $Lu = A_i \partial_i u + Bu$, we have that

$$L^* u = -Lu + (B + B^T - \partial_i A_i) u$$

Then

$$\begin{aligned}
\langle Lu, u\rangle &= -\langle u, Lu\rangle + \big\langle u, (B + B^T - \partial_i A_i) u\big\rangle \\
\langle Lu, u\rangle &= \frac{1}{2}\big\langle u, (B + B^T - \partial_i A_i) u\big\rangle \\
&\geq \frac{1}{2}\langle u, \lambda u\rangle \\
&= \frac{\lambda}{2} |u|_{L^2(Q)}^2
\end{aligned}$$

where the inequality follows from assumption. Thus the assumption tells us that L is a positive (semi)definite operator.

Now we cosnider a subspace $X \subset H \cap C^1$, $\dim X < \infty$. Let $\pi_X \colon H \to X$ be the orthogonal projection to X. Instead of solving $Lu = f$ directly, we solve $\pi_X L u = \pi_X f$, or $\pi_X(Lu - f) = 0$.

First we note that the mapping $u \mapsto \pi_X L u$ ($u \in X$) is injective, since if $\pi_X L u = 0$, then $Lu \in X^{\perp}$ and

$$0 = \langle Lu, u\rangle \geq \frac{\lambda}{2} |u|_{L^2(Q)}^2$$

for $u \in X$, and thus $u = 0$. Since X is finite dimensional, this implies that $u \mapsto \pi_x L u$ is bijective. This means that there exists a unique $u_X \in X$ such that $\pi_X L u_X = \pi_X f$.

For this u_X, we estimate $|u_X|_{L^2(Q)}$. Note that

$$\frac{\lambda}{2} |u_X|_{L^2(Q)}^2 \leq \langle L u_X, u_X\rangle = \langle \pi_X L u_X, u_X\rangle = \langle \pi_X f, u_X\rangle = \langle f, u_X\rangle \leq |f|_{L^2(Q)} |u_X|_{L^2(Q)}$$

so that $|u_X|_{L^2(Q)} \leq \frac{2}{\lambda} |f|_{L^2(Q)}$, and estimate independent of X. Here we have used the fact that orthogonal projections whose ranges are closed are self adjoint, so that $\langle \pi_X L u_X, u_X\rangle = \langle L u_X, \pi_X u_X\rangle = \langle L u_X, u_X\rangle$.

Then we take X larger and larger, closer to $C^1 \cap H$ and then take a weak limit.

Let $(\varphi_i)_{i=1}^{\infty}$ be an orthonormal basis for H, with φ_i smooth (for instance, take trigonometric polynomials), and let $X_m = \mathrm{span}\{\varphi_1, \ldots, \varphi_m\}$. Then $\pi_m L u_m = \pi_m f$ with $|u_m|_{L^2(Q)} \leq \frac{2}{\lambda} |f|_{L^2(Q)}$. u_m is then a bounded sequence of functions in H, and therefore there exists a subseqeuence $u_{m'}$ that converges to some u weakly, by the reflexivity of H and Theorem 60. Now we show that the weak limit u is the desired solution.

Note for $v \in X_{m_0}$ for $m_0 \le m$,

$$\langle u_m, L^* v\rangle = \langle L u_m, v\rangle = \langle \pi_m L u_m, v\rangle = \langle \pi_m f, v\rangle = \langle f, v\rangle$$

and taking the limit as $m \to \infty$, this shows that $\langle u, L^* v\rangle = \langle f, v\rangle$ for $v \in X_{m_0}$. Thus this is true for $v \in \bigcup_{k=1}^{\infty} X_k$. Now for any $v \in C^1 \cap H$, we can find $v_k \in \bigcup_k X_k$ with $v_k \to v$ and $D v_k \to D v$ uniformly (Weierstrass type theorem) Then

$$\langle u, L^* v\rangle = \lim_k \langle u, L^* v_k\rangle = \lim_k \langle f, v_k\rangle = \langle f, v\rangle$$

as desired. It turns out that $u \in C^1$ as well... (not proved) $\square$

Uniform Boundedness Principle

First we start with a topological theorem. Let (X, d) be a complete metric space.

Theorem 62. (Baire Category Theorem) *The intersection of a collection of dense open subsets is dense in X.*

The reason for the name is because of an immediate corollary and the definition of "first category" sets. First, we say that a set E is nowhere dense if $\left(\bar{E}\right)^c$ is dense in X. We say that E is of first category if it is the countable union of nowhere dense sets. All other sets are of second category.

Corollary 63. *Let (X, d) is a complete metric space. Then X is of second category.*

Consequently, if F_k is a sequence of closed sets with empty interior, then the union also has empty interior.

And if F_k is a sequence of closed sets whose union is X, then at least one F_k has nonempty interior.

Proof. Take complements in the Baire Category Theorem. Note a closed set is nowhere dense if and only if it does not contain an open ball. $\square$

Proposition 64. (Uniform Boundedness Principle) *Let X, Y be Banach spaces, and let A_α for $\alpha \in \Lambda$ be a family of bounded linear operators from $X \to Y$ for which*

$$\sup_{\alpha \in \Lambda} |A_\alpha(x)| < \infty$$

for all x. Then $\sup_{\alpha \in \Lambda} \|A_\alpha\| = \sup_{\alpha \in \Lambda, |x| \le 1} |A_\alpha(x)| < \infty$ (i.e. the bound is uniform)

Proof. Take $F_k = \{x \in X : |A_\alpha x| \le k\}$ for all α. Then $\bigcup_k F_k = X$ by assumption. Since F_k is closed, one of them has nonempty interior, so there exists $B_r(0) \subset F_k$. (each F_k is symmetric and convex)

Now for any $|x| < 1$, we note that $r x \in B_r(0)$, and thus

$$|A_\alpha x| = \frac{1}{r} |A_\alpha(r x)| \le \frac{k}{r}$$

for all α, and taking the supremum over $|x| < 1$, we conclude that $\|A_\alpha\| \le \frac{k}{r}$, as desired. $\square$

Applications:

Proposition 65. *If $f_i \to f$ strongly in L^p and $g_i \to g$ weakly in $L^{p'}$, then $f_i\, g_i \to f\, g$ weakly in L^1.*

Proof. We will show that $l(f_i\, g_i) \to l(f\, g)$ for $l \in (L^1)' = L^\infty$. Given $\varphi \in L^\infty$, we have that

$$\begin{aligned}
\left| \int f_i\, g_i\, \varphi - \int f\, g\, \varphi \right| &\leq \int |f_i - f|\, |g_i|\, |\varphi| + \left| \int f\, g_i\, \varphi - \int f\, \varphi\, g \right| \\
&\leq |f_i - f|_p\, |g_i|_{p'}\, |\varphi|_\infty + \left| \int f\, \varphi\, g_i - \int f\, \varphi\, g \right| \\
&\leq C\, |f_i - f|_p + \left| \int f\, \varphi\, g_i - \int f\, \varphi\, g \right| \\
&\to 0
\end{aligned}$$

where we note that $|g_i|_{p'}$ is uniformly bounded by the uniform boundedness principle. ☐

Remark 66. There exists $f_i \in L^2[0,1]$ such that $f_i \to f$ weakly in L^2, $f_i^2 \to g$ weakly in L^1 but $g \neq f^2$.

Try $f_i = \mathrm{sign}(\sin(2\pi x))$ (square wave oscillating between 1 and -1). Then $f_i \to 0$ weakly, justifying first for continuous functions and then approximating arbitrary L^2 functions by continuous functions. Also, $f_i^2 \to 1$ strongly (and hence weakly), but $g \neq f^2$.

Proposition 67. *Suppose $x_j \to x$ weakly. Then $\sup_j |x_j| < \infty$, i.e. the sequence is bounded.*

Proof. Recall the mapping κ: $X \to X''$ given by $x \mapsto \kappa x$ where $\kappa x(l) = l(x)$, which is an isometric mapping, so that $|x|_X = |\kappa x|_{X''}$. Then $\kappa x_j \in X''$ and given $l \in X'$,

$$\sup_j |\kappa x_j(l)| = \sup_j |l(x_j)| < \infty$$

since $l(x_j)$ converges. Thus the conditions in Proposition 64 are satisfied, and

$$\sup_j |x_j|_X = \sup_j |\kappa x_j|_{X''} < \infty$$

as desired. ☐

Lower Semicontinuity under Weak Convergence

Remark 68. If K is closed with respect to the norm, and $x_j \in K$, $x_j \to x$ weakly. It is not necessarily true that $x \in K$. A quick example is for instance $K = \{e_j, j \in \mathbb{N}\} \subset l^2$. Then $e_j \to 0$ weakly since for any $(x_k)_{k=1}^\infty \in l^2$,

$$\lim_{j\to\infty} \langle e_j, (x_k)_{k=1}^\infty \rangle = \lim_{j\to\infty} x_j = 0$$

but $0 \notin K$. This says that the weak-closure of a set is not necessarily the same as the norm-closure.

However, if we add the condition that the set is convex, then the property does hold. That is, the weak-closure of a convex set is the same as the norm-closure.

Proposition 69. *Let K is convex and norm-closed, and suppose that $x_j \in K$ and $x_j \to x$ weakly. Then $x \in K$.*

Proof. Suppose that $x \notin K$. Then there exists $r > 0$ such that $B_r(x) \cap K = \varnothing$. Then there exists $l \in X'$ such that $l \neq 0$ and $\operatorname{Re} l|_K \leq \operatorname{Re} l|_{B_r(x)}$. Then using y such that $\operatorname{Re} l(y) = 1$ (pick arbitrarily and scale appropriately), we have that

$$\operatorname{Re} l|_K \leq \operatorname{Re} l(x - \varepsilon y) = \operatorname{Re} l(x) - \varepsilon$$

for $\varepsilon < r$. However, this implies that $\operatorname{Re} l(x_j) \leq \operatorname{Re} l(x) - \varepsilon$, and taking limits, $\operatorname{Re} l(x) \leq \operatorname{Re} l(x) - \varepsilon$, a contradiction. □

Corollary 70. *If $f: X \to \mathbb{R}$ convex, continuous, and $x_j \to x$ weakly, then*

$$f(x) \leq \liminf_j f(x_j)$$

Proof. Suppose that $c = \liminf_{j \to \infty} f(x_j)$. Consider the set $K = \{x: f(x) \leq c + \varepsilon\}$. This is a (norm)-closed, convex set by assumption. By definition of liminf, we can find a subsequence $x_{j'}$ such that $f(x_{j'}) < c + \varepsilon$. Since $x_{j'} \to x$ weakly and K is convex and closed, Proposition 69 shows that $x \in K$ so that $f(x) \leq c + \varepsilon$. Since ε is arbitrary, we have proved the result. □

Example 71. Let X be reflexive, $K \subset X$ be closed, convex, and nonempty. Suppose $y \notin K$. Then there exists $x_0 \in K$ such that $|y - x_0| = \operatorname{dist}(y, K)$.

Proof. Recall $\operatorname{dist}(y, K) = \inf_{x \in K} |y - x|$. By definition of dist, there exists x_n such that $|y - x_n| \to \operatorname{dist}(y, K)$. Then since $|x_n| \leq |y - x_n| + |y|$, we note that x_n is a bounded sequence in a reflexive space. Thus by Theorem 60 there exists a subsequene $x_{j'}$ that converges weakly to some x_0. Since K is closed and convex, $x_0 \in K$. Then by Corollary 70, and the fact that the mapping $x \mapsto |y - x|$ is continous and convex,

$$|y - x_0| \leq \liminf_j |y - x_{j'}| = \operatorname{dist}(y, K)$$

and since $\operatorname{dist}(y, K) \leq |y - x_0|$ by definition of dist, so that $|y - x_0| = \operatorname{dist}(y, K)$. □

Upgrading to Norm Convergence

Recall that a sequence that converges in norm also converges weakly, and that the converse is not true. However, with an additional condition, it is true.

Proposition 72. *If $x_j \to x$ weakly, and $|x_j| \to |x|$, and X is uniformly convex, then $x_j \to x$ in norm.*

In particular, this is true for reflexive spaces such as L^p for $1 < p < \infty$.

Proof. If $x = 0$, then there is nothing to prove.

Suppose that $x \neq 0$. Then consider $u_j = \frac{x_j}{|x_j|}$ and $u = \frac{x}{|x|}$. Note $u_j \to u$ weakly since

$$l(u_j) = \frac{l(x_j)}{|x_j|} \to \frac{l(x)}{|x|} = l(u)$$

Furthermore, $\frac{u_j + u_k}{2} \to u$ weakly as $j, k \to \infty$. By lower semicontinuity of the norm,

$$1 = |u| \leq \liminf_{j,k} \left| \frac{u_j + u_k}{2} \right| \leq \limsup_{j,k} \left| \frac{u_j + u_k}{2} \right| \leq 1$$

This implies that $\left|\frac{u_j+u_k}{2}\right| \to 1$ and by uniform convexity, $|u_j - u_k| \to 0$. Therefore u_j is Cauchy in norm and $u_j \to u$ in norm (recall the weak limit is unique, so u_j must converge to u). This means that

$$|x_j - x| = ||x_j|\, u_j - |x|\, u| \le |x_j|\, |u_j - u| + ||x_j| - |x||\, |u| \to 0$$

□

Remark 73. To see what can happen in $L^1[0,1]$ say, which is not uniformly convex (not even strictly convex), consider $\varphi_n = \mathrm{sign}(\sin(2\pi n x))$ (square waves). Then since $\varphi_n \to 0$ weakly, $1+\varphi_n \to 1$ weakly. But $|1+\varphi_n|_1 = 1$ since φ_n is $\{1,-1\}$ each on a set of measure $1/2$, thus $1+\varphi_n$ is $\{0,1\}$ each on a set of measure $1/2$. Thus $1+\varphi_n \to 1$ weakly, $|1+\varphi_n|_1 \to |1|_1$, but $|1+\varphi_n - 1|_1 = |\varphi_n|_1 \not\to 0$, so there is no norm convergence.

(condition for L^1... possibly involving the additional assumption that $f_n \to f$ in measure for all finite sets. Recall, norm convergence in L^1 implies that $f_n \to f$ in measure on finite subsets. The converse is true if we add weak convergence.)

Weak* Convergence

Let X be a Banach space, and $l_j \in X'$. We say that l_j converges to l in the **weak*** sense if $l_j(x) \to l(x)$ for all $x \in X$.

Theorem 74. *If X is separable, then every bounded sequence $l_j \in X'$, $|l_j| \le c$ has a subsequence $l_{j'}$ that converges in the weak* sense.*

Example 75. Let X be a compact metric space, and (X,μ) with μ a Borel measure. Then any function in $L^1(X,\mu)$ can be considered a functional in $C(X)'$ by considering $f \mapsto l_f = f d\mu$. In other words, given a continuous function $\varphi \in C(X)$, $l_f(\varphi) = \int \varphi\, d\mu$. Then, given a sequence $f_j \in L^1$ that is bounded $|f_j|_{L^1} \le c$, there exists some measure ν on X such that $f_j\, d\mu \to \nu$ in the weak* sense (or a subsequence does), i.e.

$$\int_X \varphi\, f_j\, d\mu \to \int_X \varphi\, d\nu$$

for all $\varphi \in C(X)$.

Remark 76. Proposition 64 (Uniform Boundedness Principle) implies that if $l_j \to l$ in the weak* sense, then $\sup_j |l_j| < \infty$. In the hypotheses of the proposition, $Y = \mathbb{K}$, and for $x \in X$, $\sup_j l_j(x) < \infty$.

Example 77. *Quadratures for Integral Evaluation*. Suppose we want to find a suitable rule for approximating $\int_0^1 f(x)\, dx$ for continuous f using a family of weighted sums $\sum_{k=1}^n f(x_j^{(n)})\, w_j^{(n)}$. Note that a necessary condition for

$$\lim_{n\to\infty} \sum_{k=1}^n f(x_j^{(n)})\, w_j^{(n)} = \int_0^1 f(x)\, dx$$

for all f is the uniform boundedness of the functionals $l_n(f) = \sum_{k=1}^n f(x_j^{(n)})\, w_j^{(n)}$, since weak*-ly convergent sequences are bounded. Defining $l(f) = \int_0^1 f(x)\, dx$, we are saying that if $l_n \to l$ in the weak* sense, then l_n is a bounded sequence.

Remark 78. Note that on X' we can also talk about weak convergence. That is, if $l_n \in X'$, $l \in X'$, then $l_n \to l$ weakly if for all $\varphi \in X''$, $\varphi(l_n) \to \varphi(l)$.

Because $\kappa \colon X \to X''$ is an isometric embedding, if $l_n \to l$ weakly, then $l_n \to l$ in the weak* sense. To see this, for any $x \in X$, we have that for κx (defined by $\kappa x(l) = l(x)$), $\kappa x(l_n) \to \kappa x(l)$, and thus $l_n(x) \to l(x)$.

The converse is not true in general, but it holds when X is reflexive. This is because $\kappa(X) = X''$, so that if $l_n \to l$ in the weak* sense, then given $\varphi \in X''$, we can find $x \in X$ so that $\varphi(l) = l(x)$ for all l, and thus $\varphi(l_n) \to \varphi(l)$ for all φ, and $l_n \to l$ weakly.

Weak and Weak* Convergence in L^1

Example 79. Consider $C[0,1]'$, which is the space of regular Borel measures on $[0,1]$, where

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

and $|\mu|_{C[0,1]'} = (\mu_+ + \mu_-)([0,1]) = |\mu|([0,1])$ ($|\mu|$ is the total variation measure).

Note that $L^1[0,1]$ can be isometrically embedded into $C[0,1]'$ by identifying f with $f dm$ where m is the Lebesgue measure. Then consider $f_n = n\,\mathbf{1}_{[0,1/n]} \in L^1[0,1]$. Then for every $\varphi \in C[0,1]$, we have that

$$\int f_n \varphi \to \varphi(0) = \int \varphi \, d\delta_0$$

where δ_0 is the dirac delta measure $\delta_0(A) = \left\{ \begin{smallmatrix} 1 & 0 \in A \\ 0 & 0 \notin A \end{smallmatrix} \right.$. Thus $f_n \, dm \to \delta_0$ in the weak* sense. However, f_n does not converge to any $f \in L^1$ in the (L^1)-weak sense (recall $(L^1)' \cong L^\infty$).

Proof. Suppose towards a contradiction that f_n does converge weakly to some function $f \in L^1$. Using $\mathbf{1}_{[0,t]} \in L^\infty$, weak convergence of f_n shows that

$$1 = \lim_{n \to \infty} \int_0^t f_n = \int_0^t f$$

Then letting $F(t) = \int_0^t f$, we have that $F(t) = 1$, and because f is integrable, by the Fundamental Theorem of Calculus $F' = f$. But $F' = 0$ (since F is constant), and thus $f = 0$, which contradicts $\int_0^t f = 1$. $\square$

Moreover, $f_n \, dm$ does not converge to δ_0 in the ($C[0,1]'$)-weak sense either!

Proof. Denote $M = C[0,1]'$, the space of Borel measures. Note that $M' \subset L^\infty$ since $M \supset L^1$. Also, for any measurable set B, we can define a linear functional l_B on M' by $l_B(\mu) = \mu(B)$. Note l_B is linear and bounded since $|l_B(\mu)| = |\mu(B)| \leq |\mu|([0,1]) \leq 1$ for $|\mu| \leq 1$. Thus, using $(0,t]$, we have that

$$l_{(0,t]}(f_n \, dm) = \int_0^t f_n \, dm \to 1$$

But

$$l_{(0,t]}(\delta_0) = 0$$

Thus $f_n \, dm$ does not converge to δ_0 weakly. (Note that if $f_n \, dm$ converges weakly to some ν, then it converges to ν in the weak* sense, and since the weak* limit is unique $\nu = \delta_0$. Thus, it cannot converge to some other $\nu \neq \delta_0$ weakly) $\square$

This situation also provides an example of a bounded sequence in L^1 with no weakly-convergent subsequence.

When L^1 functions have a weak limit in L^1

A natural followup question is then, when does a sequence $f_i \in L^1$ have a weakly-convergent subsequence in L^1? We already know that boundedness is not enough to have a weakly-convergent subsequences. It turns out it is enough to add the property that the functions are uniformly integrable.

Proposition 80. *Let $(X,\, \mu)$ be a finite measure space where $L^1(X)$ is separable. Let $f_j \in L^1(X)$ be a bounded sequence, i.e. $\sup |f_j|_{L^1} < \infty$, and furthermore, suppose that f_j satisfy*

- *(Uniform Integrability) For any $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mu(A) < \delta$, then*

  $$\int_A |f_j| d\mu \leq \varepsilon$$

  for all j.

Then there exists a subsequence $f_{j'}$ converging weakly to some $f \in L^1$.

Proof. The idea is to show that for measurable A, $\int_A f_{j'}\, d\mu \to \nu(A)$ for some measure ν and subsequence $f_{j'}$ and then that $\nu \ll \mu$.

Step 1. First we find a countable collection of sets A_i for which we can approximate an arbitrary measurable set A; in other words, given $\varepsilon > 0$, we can find A_i such that $\mu(A \vartriangle A_i) < \varepsilon$. To do this, consider the set $S = \{\mathbf{1}_A \colon A \text{ measurable}\} \subset L^1(X)$. This is a closed subset, since given a sequence $\mathbf{1}_{A_j}$ for which $\mathbf{1}_{A_j} \to f$ in L^1,

$$\mu(\{|f| > \varepsilon\} \cap \{|f - 1| > \varepsilon\}) \leq \mu(\{|f - \mathbf{1}_{A_j}| > \varepsilon\}) \leq \frac{1}{\varepsilon} \int |f - \mathbf{1}_{A_j}| \to 0$$

so that $\mu(\{f \notin \{0, 1\}\}) = 0$, and $f = \mathbf{1}_A$ for some A. Thus S is closed, and since L^1 is assumed to be separable, S is also separable, and there exists a countable collection $\{\mathbf{1}_{A_i},\ i \in \mathbb{N}\} \subset S$ which is dense in S. This means given an arbitrary measurable A, and $\varepsilon > 0$, there exists A_i such that

$$\mu(A_i \vartriangle A) = \int |\mathbf{1}_{A_i} - \mathbf{1}_A| d\mu < \varepsilon$$

Step 2. The next step is to extract a subsequence of f_j for which $\int_{A_i} f_{j'} d\mu$ converges for all A_i. This is accomplished via a diagonalization process. Since $\int_{A_1} f_j\, d\mu$ is a bounded sequence,

$$\sup_j \left| \int_{A_1} f_j\, d\mu \right| \leq \sup_j |f_j|_{L^1} < \infty$$

there is a subsequence $f_{1,j}$ for which $\int_{A_1} f_{1,j}\, d\mu$ converges. Likewise, there is a subsequence $f_{2,j}$ of $f_{1,j}$ for which $\int_{A_2} f_{2,j}\, d\mu$ converges. The key is that since $f_{2,j}$ is a subsequence of $f_{1,j}$, we also have that $\int_{A_1} f_{2,j}\, d\mu$ converges. Continuing onwards, we get nested subsequences $\{f_{1,j}\} \supset \{f_{2,j}\} \supset \ldots$ and if we take the sequence $(f_{1,1}, f_{2,2}, f_{3,3}, f_{4,4}, \ldots)$ this gives the desired sequence for which we can define

$$\nu(A_i) = \lim_{j \to \infty} \int_{A_i} f_{j,j}\, d\mu$$

Step 3. The third step is to show that $\int_A f_{j,j}\,d\mu$ converges to some $\nu(A)$ using density of A_i and the uniform integrability of $f_{j,j}$. To show this, given $\varepsilon > 0$, let δ be as in the uniform integrability property in the hypotheses of the theorem. Choose A_i such that $\mu(A \triangle A_i) < \delta$, and let $N > 0$ such that for $j, k > N$, $\left|\int_{A_i} f_j\,d\mu - \int_{A_i} f_k\,d\mu\right| < \varepsilon$ (note $\int_{A_i} f_j\,d\mu$ converges, and thus is Cauchy). Then for $j, k > N$,

$$\begin{aligned}\left|\int_A f_j\,d\mu - \int_A f_k\,d\mu\right| &\le \left|\int_A f_j\,d\mu - \int_{A_i} f_j\,d\mu\right| + \left|\int_{A_i} f_j\,d\mu - \int_{A_i} f_k\,d\mu\right| + \left|\int_{A_i} f_k\,d\mu - \int_A f_k\,d\mu\right| \\ &\le \int_{A \triangle A_i} |f_j|\,d\mu + \left|\int_{A_i} f_j\,d\mu - \int_{A_i} f_k\,d\mu\right| + \int_{A \triangle A_i} |f_k|\,d\mu \\ &\le 3\varepsilon\end{aligned}$$

Thus $\int_A f_j\,d\mu$ is a Cauchy sequence, and we set the value of $\nu(A)$ to be the limit.

Step 4. The fourth step is to show that ν is a (signed) measure with $\nu \ll \mu$. Finite additivity follows from linearity of limits and integrals. For countable additivity, note that $\nu(X) < \infty$, and given a nested sequence of sets $A_j \supset A_{j+1}$ converging to the empty set $\varnothing$, we see that $|\nu(A_j)| \to 0$ by uniform integrability of f_j. Also, if $\mu(A) = 0$, then $\nu(A) = \lim_j \int_A f_j\,d\mu = 0$, so that $\nu \ll \mu$. Thus by Radon-Nikodym, there exists $f \in L^1$ for which $\nu = f\,d\mu$.

Step 5. Finally, we have found $f \in L^1$ for which

$$\int_A f_j\,d\mu \to \int_A f\,d\mu$$

for all measurable sets A. By linearity we have that $\int_A f_j\,\varphi\,d\mu \to \int_A f\,\varphi\,d\mu$ for simple functions φ, and by density we conclude that this holds for all $\varphi \in L^\infty$, and this shows that $f_j \to f$ weakly, as desired.

□

The converse is also true, that the weak convergence of f_j in L^1 implies the uniform integrability of f_j. This follows as a corollary to the following theorem:

Theorem 81. (Vitali-Hahn-Saks) *Assume (X, μ) is a finite measure space, and ν_i is a sequence of signed measures on X with $\nu_i \ll \mu$ and ν_i converges setwise to some set function ν, i.e. $\nu_i(A) \to \nu(A)$ for all measurable A. Then:*

1. *For all $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mu(A) < \delta$, then $|\nu_i|(A) < \varepsilon$ for all i. (this is absolute continuity, except uniform in i)*

2. *ν is also a signed measure*

Proof. Let $Y = \{A \subset X, A\,\text{measurable}\}$ under the metric $d(A, B) = |\mathbf{1}_A - \mathbf{1}_B|_{L^1} = \mu(A \triangle B)$. Then (Y, d) is a complete metric space (we proved that Y is a closed subspace of L^1 in Proposition 80). With this metric, note that $\nu_i: Y \to \mathbb{C}$ is continuous. This is because $\nu_i \ll \mu$ so that $\nu_i = f_i\,d\mu$ for some $f_i \in L^1$, $f_i \ge 0$. Then

$$|\nu_i(A) - \nu_i(B)| = \nu_i(A \triangle B) = \int_{A \triangle B} f_i\,d\mu$$

and by absolute continuity of the integral, if $d(A, B) = \mu(A \triangle B)$ is small, we can make the RHS small as well. In particular, if $B = \varnothing$, in the context of property (1) above, given ε we can find a δ_i that works for individual i, and also for finitely many i (by taking the minimal δ_i). It remains to study what happens as $i \to \infty$.

Now let $\varepsilon > 0$. Define

$$F_n = \{A \in Y \colon |\nu_k(A) - \nu_n(A)| \leq \varepsilon, k \geq n\}$$

F_n is a nested sequence of closed subspaces increasing to Y. By the corollary to Baire's Category Theorem (Corollary 63), one of the F_n has non-empty interior. So we can find $B_r(A_0) \subset F_n$ for some n, A_0, r. This means that for any set $B \in B_r(A_0)$, $|\nu_k(B) - \nu_n(B)| \leq \varepsilon$.

Then note that for any set A with $\mu(A) < r$, we can write $A = (A \cup A_0) \backslash (A_0 \backslash A)$ where $A \cup A_0$ and $A_0 \backslash A$ are both in $B_r(A_0)$:

$$d(A \cup A_0, A_0) = \mu((A \cup A_0) \triangle A_0) = \mu(A \backslash A_0) \leq \mu(A) < r$$

$$d(A_0 \backslash A, A_0) = \mu((A_0 \backslash A) \triangle A_0) = \mu(A \cap A_0) \leq \mu(A) < r$$

Thus for $k \geq n$, we have that

$$\begin{aligned}
|\nu_k(A) - \nu_n(A)| &= |\nu_k(A \cup A_0) - \nu_k(A_0 \backslash A) + \nu_n(A \cup A_0) - \nu_n(A_0 \backslash A)| \\
&\leq |\nu_k(A \cup A_0) - \nu_n(A \cup A_0)| + |\nu_k(A_0 \backslash A) - \nu_n(A_0 \backslash A)| \\
&\leq 2\varepsilon
\end{aligned}$$

To finish the result, by the absolute continuity of ν_n, we can find a δ_n for which $\mu(A) < \delta_n \Longrightarrow \nu_n(A) < \varepsilon$. Then for $k \geq n$, this same δ_n shows that for $\mu(A) < \delta_n$,

$$|\nu_k(A)| \leq |\nu_k(A) - \nu_n(A)| + |\nu_n(A)| \leq 3\varepsilon$$

and so we have found a δ_n in property (1) that works for all $k \geq n$. Combining this with the observation that we can already find $\delta_1, \ldots, \delta_{n-1}$ that works for $1, \ldots, n-1$, choosing $\delta = \min(\delta_1, \ldots, \delta_n)$ proves property (1).

Showing that ν is a measure is straightforward, and is proved in the same manner as in step 4 in the proof of Proposition 80. $\square$

Corollary 82. *Let (X, μ) be a finite measure space. Let f_i be a sequence of L^1 functions converging to $f \in L^1$ weakly. Then f_i is uniformly integrable, i.e. for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all A with $\mu(A) < \delta$, $\int_A |f_i| d\mu < \varepsilon$.*

Proof. Note if $f_i \to f$ weakly, then using $\varphi = \mathbf{1}_A$ gives $\int_A f_i \, d\mu \to \int_A f d\mu$. Since $f_i \, d\mu$ is a sequence of signed measures converging to a set function $f \, d\mu$ (which is already a measure), property (1) of Theorem 81 gives the result. Note $|f_i \, d\mu|(A) = \int_A |f_i| \, d\mu$. $\square$

Upgrading Weak Convergence to Norm Convergence

Recall Proposition 72, about when weak convergence implies convergence in norm. For L^1, we have the following result:

Theorem 83. *Let (X, μ) be a finite measure space. Suppose $f_n, f \in L^1(X, \mu)$ and $f_n \to f$ weakly. Then if in addition, $f_n \to f$ in measure, then $f_n \to f$ in norm.*

Proof. The idea for this particular proof is to use the following real analysis lemma (there is probably a better way):

Lemma 84. *Let (X,μ) be a finite measure space. Suppose $f_n, f\in L^1(X)$ and $f_n\to f$ a.e. Then*

$$|f_n-f|_{L^1}\to 0 \Longleftrightarrow |f_n|_{L^1}\to|f|_{L^1}$$

Proof. ($\Longrightarrow$) This direction is obvious by triangle inequality $\left||f_n|_{L^1}-|f|_{L^1}\right|\leq|f_n-f|_{L^1}$

($\Longleftarrow$) This direction is by generalized Dominated Convergence Theorem. Note that we just need to justify the limit swap in

$$\lim_n\int|f_n-f|=\int\lim_n|f_n-f|=0$$

Note that $|f_n-f|\leq|f_n|+|f|$ and $\int|f_n|+|f|\to\int 2|f|$, i.e. $|f_n-f|$ is dominated by an integrable sequence g_n for which $\int g_n\to\int\lim_n g_n$. In addition, $f_n\to f$ a.e., and we have satisfied all conditions in the generalized Dominated Convergence Theorem (e.g. see Royden). Thus the limit swap is valid, and we have proved the result. □

Note that although we need $f_n\to f$ a.e. to use the Lemma, if $f_n\to f$ in measure, there is a subsequence $f_{n'}$ for which $f_{n'}\to f$ a.e. Thus, the goal is to show that for this subsequence, $|f_{n'}|_{L^1}\to|f|_{L^1}$. For the rest of the prove, we will use f_n to denote the *subsequence* that converges a.e.

Let $\varepsilon>0$, and let $E_n=\{|f_n-f|>\varepsilon\}$. Then convergence in measure means $\mu(E_n)\to 0$. Also, since $f_n\to f$ weakly, the corollary to Vitali-Hahn-Saks (Corollary 82) implies that f_n is uniformly integrable. That is, we can find δ so that $\mu(A)<\delta\Rightarrow\int_A|f_n|<\varepsilon$.

Weak convergence (lower semicontinuity of norm) also implies that $|f|_{L^1}\leq\liminf_n|f_n|_{L^1}$. Now for n sufficiently large, $\mu(E_n)<\delta$, and in this case

$$\begin{aligned}\int|f_n| &= \int_{E_n}|f_n|+\int_{E_n^c}|f_n|\\ &\leq \varepsilon+\int_{E_n^c}|f|+\varepsilon\\ &\leq \int|f|+2\varepsilon\end{aligned}$$

where in the second line, we have used uniform integrability and the fact that $|f_n|\leq|f|+|f_n-f|\leq|f|+\varepsilon$ on E_n^c. Now taking the limsup, we have that

$$\limsup_n|f_n|_{L^1}\leq|f|_{L^1}+2\varepsilon$$

and since ε is arbitrary, we have that

$$|f|_{L^1}\leq\liminf_n|f_n|_{L^1}\leq\limsup_n|f_n|_{L^1}\leq|f|_{L^1}$$

or in other words, $|f_n|_{L^1}\to|f|_{L^1}$. Then by the Lemma, this implies that $|f_n-f|_{L^1}\to 0$, as desired.

What we have shown is that given the hypotheses of this Theorem, we can find a subsequence for which $f_{n'}\to f$ in norm. A technical point is that this implies that the original sequence $f_n\to f$ in norm. This is due to the fact that if every subsequence of f_n has a further subseqeuence that converges to f in norm, then f_n converges to f in norm. □

Remark 85. The proof above can be extended to σ-finite measure spaces, but changing $f_n\to f$ in measure to $f_n|_A\to f|_A$ in measure (with respect to $\mu|_A$) for all finite measure subsets $A\subset X$. Then it can be proved in the usual manner by splitting X into a countable union of (disjoint) finite measure sets and applying the above results to each individual piece.

Corollary 86. *In l^1, weak convergence is the same as norm convergence.*

Proof. Recall $l^1 = L^1(\mathbb{N}, \nu)$ with ν being the counting measure, so that convergence in measure (on finite subsets) is the same as weak convergence. Then apply the previous theorem.

Sketch of Alternate Proof: (do not need all the machinery above). It suffices to show that if $x_n \to 0$ weakly, then $|x_n|_{l^1} \to 0$. Assume towards a contradiction that $|x_n|_{l^1} \nrightarrow 0$. Then since x_n converges weakly, $|x_n|_{l^1}$ is uniformly bounded, and thus there exists a subsequence for which $|x_n|_{l^1} \approx C > 0$ for all n. Another consequence of weak convergence is that $x_n \to 0$ pointwise (use functionals $\varphi = e_k$). The idea is to find a subsequence x_{k_n} for which the essential supports are disjoint (where the l^1 norm is concentrated). Here is the construction:

Let x_{k_1} be arbitrary, and since $x_{k_1} \in l^1$ there exists some set $A_1 = [0, n_1]$ outside of which x_{k_1} has very small l^1 norm.

Now since $x_n \to 0$ pointwise, we can find x_{k_2} which almost vanishes in A_1 (a finite set). Then there exists some set $A_2 = [n_1 + 1, n_2]$ outside of which x_{k_2} has small l^1 norm.

Continuing in this manner, we have extracted a subsequence x_{k_n} for which x_{k_n} is essentially supported on A_n, where A_n are disjoint. Now let $\varphi(i)_{i=0}^{\infty}$ be the sequence defined by $\varphi(i) = \operatorname{sgn}(x_{k_n})$ when $i \in A_n$ (matches the sign of φ_{k_n} where φ_{k_n} is essentially supported). Then $\sum_i x_{k_n}(i)\varphi(i) \approx |x_{k_n}|_{l^1} \approx C > 0$ for all n, which contradicts weak convergence (which says that $\sum_i x_{k_n}(i)\varphi(i) \to 0$ as $n \to \infty$). ☐

Summary of Results: Let (X, μ) be a finite measure space.

- For $1 < p < \infty$, bounded sequences in L^p have weakly convergent subsequences.
- For $p = 1$,
  - Let X be compact, Hausdorff, and μ be a regular Borel measure. Then if we consider L^1 as a subspace of $C(X, \mathbb{C})'$ with the (norm-preserving) identification $f \mapsto f d\mu$, then a bounded sequence in L^1 has a subsequence that converges in the weak* sense to some measure $\nu \in C(X, \mathbb{C})'$.
  - A bounded sequence in L^1 has a weakly convergent subsequence if and only if the sequence is uniformly integrable.
  - Suppose a sequence in L^1 converges weakly. Then it also converges in norm if and only if it converges in measure.

Open Mapping Theorem

Theorem 87. *Let X, Y be Banach spaces, $A \in \mathcal{L}(X, Y)$ such that $A\,X = Y$, then A is an open map (maps open sets to open sets)*

Proof. To show A is an open map, we only need to show that $A(B_1) \supset B_\delta$ for some $\delta > 0$. The reason is that for arbitrary open sets U, given $y \in A(U)$, there exists some x such that $Ax = y$. Since U is open, we can find $x + B_r \subset U$. But then we can find $y + B_{\delta'} \subset A(B_r(x)) \subset A(U)$ so that $A(U)$ is open.

Now note $Y = \bigcup_{k=1}^{\infty} \overline{A(B_k)}$, and by the corollary to Baire Category Theorem (Corollary 63), one of the $\overline{A(B_k)}$ has nonempty interior, so that $B_r(y) \subset \overline{A(B_k)}$ for some $r, k > 0, y \in \overline{A(B_k)}$

- First show we can find $B_r(0) \subset \overline{A(B_1)}$ for some different r. (Since $\overline{A(B_k)}$ is convex, symmetric and contains 0, we should expect this to be true). Note that

$$B_{r/2}(0) \subset B_{r/2}(y) - B_{r/2}(y) \subset B_r(y) \subset \overline{A(B_k)}$$

  and since $\overline{A(B_k)} = k\overline{A(B_1)}$, we note that $B_{r/2k}(0) \subset \overline{A(B_1)}$.

- Now we show that $B_r(0) \subset A(B_2)$ given above. Let $|x| < r$. We want to find $|z| < 2$ such that $A\,z = x$.

  Since $B_r(0) \subset \overline{A(B_1)}$, we can find $|z_0| < 1$ with $|x - A\,z_0| < \frac{r}{2}$.

  Then $x - A\,z_0 \in B_{r/2}$, and since $B_{r/2}(0) \subset \overline{A(B_{1/2})}$, we can find $|z_1| < \frac{1}{2}$ with

$$|x - A\,z_0 - A\,z_1| < \frac{r}{4}$$

  Continuing in this manner, we have a sequence $|z_k| < \frac{1}{2^k}$ with

$$\left| x - \sum_{j=0}^{k} A\,z_k \right| < \frac{r}{2^k}$$

  Since $\sum_k |z_k| < \infty$, $\sum_k z_k$ is Cauchy and converges to some z, and then

$$|x - A\,z| = \lim_k \left| x - \sum_{j=0}^{k} A\,z_k \right| = 0$$

  with $|z| < \sum_k |z_k| = 2$ and $A\,z = x$.

  We conclude that $B_{r/2}(0) \subset A(B_1)$ as desired.

□

This has many important consequences.

Example 88. If X, Y Banach, $A \in \mathcal{L}(X, Y)$ is bijective, then A^{-1} is also continuous, and there exists $c > 0$ for which

$$c^{-1}|x| \leq |Ax| \leq c|x|$$

Since A is an open map, the preimage of A^{-1} of any open set is open, and thus A^{-1} is continuous.

Theorem 89. (Closed Graph Theorem) *Let X, Y be Banach spaces, and let A: $X \rightarrow Y$ be a linear map. Let the graph of A, G_A be defined by*

$$G_A = \{(x, A\,x), x \in X\} \subset X \times Y$$

where $X \times Y$ is also a Banach space. Then if G_A is closed, then A is bounded.

Remark 90. Note that if A is bounded, then G_A is closed, since $(x_j, A\,x_j) \to (x, y)$ implies that $x_j \to x$ and $A\,x_j \to y$. But since A is bounded, and hence continuous, $A\,x_j \to A\,x$ so $A\,x = y$.

Proof. Consider the projection operator $\pi: G_A \to X$ defined by $\pi(x, A\,x) = x$. π is continuous and since G_A is closed π is a bijection between two Banach spaces G_A and X. Thus π is an open map, and we can find c such that

$$c^{-1}|(x, A\,x)|_{X\times Y} \le |\pi(x, A\,x)| = |x|$$

and

$$c^{-1}(|x| + |A\,x|) \le |x|$$

so that

$$|A\,x| \le (c+1)|x|$$

and thus A is bounded. □

Example 91. Let H be a Hilbert space over $\mathbb{R}$, and let $A: H \to H$ be a linear operator such that there exists $B: H \to H$ linear with $\langle A\,x, y\rangle = \langle x, B\,y\rangle$, then A is bounded.

To prove, we show that G_A is closed. Let $(x_k, A\,x_k) \to (x, y)$. We then want to show that $y = A\,x$. Note that $x_k \to x$ and $A\,x_k \to y$, and then for any $z \in H$, we have

$$\langle A\,x_k, z\rangle = \langle x_k, B\,z\rangle$$

we take limits to obtain

$$\langle y, z\rangle = \langle x, B\,z\rangle = \langle A\,x, z\rangle$$

for all z. Then $\langle y - A\,x, z\rangle = 0$ and taking $z = y - A\,x$ shows that $y = A\,x$. This implies that A is bounded.

Example 92. Let X be a Banach space. Let $X = X_1 \oplus X_2$ with X_1, X_2 closed. Take $\pi_1: X \to X_1$ and $\pi_2: X \to X_2$. Then π_1, π_2 are continuous. That is, if we decompose $x = x_1 + x_2$ with $x_i \in X_i$, then

$$|x_1| \le c\,|x| \text{ and } |x_2| \le c\,|x|$$

for some constant c.

Note if we define a new norm on X by $\|x\| := |\pi_1\,x| + |\pi_2\,x|$, then π_1 and π_2 are continuous with respect to $\|\cdot\|$, since $\|\pi_1 x\| + \|\pi_2 x\| = |x_1| + |x_2| = \|x\|$, so individually $\|\pi_i x\| \le \|x\|$. Then it suffices to show that the identity map from $(X, \|\cdot\|) \to (X, |\cdot|)$ is a homeomorphism.

Note that $(X, \|\cdot\|)$ is complete. Suppose that $x^{(j)}$ is Cauchy in $(X, \|\cdot\|)$ so that $\|x^{(j)} - x^{(k)}\| \to 0$. Then individually, $|x_i^{(j)} - x_i^{(k)}| \to 0$ for $i \in \{1, 2\}$. Then by completeness of X_1, X_2 as closed subspaces of $(X, |\cdot|)$, there exists x_i for which $|x_i^{(j)} - x_i| \to 0$. But then $\|x^{(j)} - (x_1 + x_2)\| \to 0$, so that $x^{(j)}$ converges to $x_1 + x_2$ in $\|\cdot\|$.

Then the identity map between two complete spaces is a homeomorphism by the open mapping theorem, and thus there exists c for which $c^{-1}\|x\| \le |x|$. This means that

$$|x_i| \le |x_1| + |x_2| \le c\,|x|$$

for $i \in \{1, 2\}$, as desired.

Example 93. Consider $l^\infty \supset c_0 = \{x_k: x_k \to 0 \text{ as } k \to \infty\}$. The claim is that c_0 has no closed complement. Thus it is not necessarily true that a closed subspace has a closed complement.

Suppose towards a contradiction that $l^\infty = c_0 \oplus Z$ for some Z closed. Then by the previous example, the projection $\pi: l^\infty \to c_0$ must be continuous.

First we claim that for $\varphi_i \in (l^\infty)'$, if $\varphi_i \to 0$ weak*-ly, then

$$\lim_{i \to \infty} \sum_{k=1}^{\infty} |\varphi_i(e_k)| = 0$$

where $e_k(i) = \left\{ \begin{smallmatrix} 1 & i = k \\ 0 & i \neq k \end{smallmatrix} \right.$ (standard basis).

Given the claim for now, we define $\psi_i: c_0 \to \mathbb{C}$ by $\psi_i(x) = x_i$, and note that $\psi_i \to 0$ weak*-ly, since given any $x \in c_0$, $\psi_i(x) = x_i \to 0$ by definition of c_0. Note $|\psi_i(x)| = |x_i| \leq |x|_{l^\infty}$ so that ψ_i are bounded, and thus in $(c_0)'$. Then if we define $\varphi_i = \psi_i \circ \pi$, we have $\varphi_i: l^\infty \to \mathbb{C}$ and $\varphi_i \to 0$ weak*-ly as well. Also, as φ_i is a composition of two bounded operators, $\varphi_i \in (l^\infty)'$.

However, $\sum_{k=1}^{\infty} \varphi_i(e_k) = 1$ since $\varphi_i(e_k) = \psi_i(e_k) = \delta_{ik}$, which contradicts the claim.

To prove the claim, we need to prove that $a_i = (\varphi_i(e_k))_{k=1}^{\infty}$ is a sequence in l^1 which converges to 0 strongly. We will use the result that in l^1, weak convergence is the same as norm convergence (Corollary 86), and show that $a_i \to 0$ weakly in l^1. First of all, each $a_i \in l^1$ by applying φ_i to $b_k = \operatorname{sgn} a_i(k) \in l^\infty$

$$\varphi_i(b_k) = \sum_{k=1}^{\infty} \operatorname{sgn} a_i(k)\, \varphi_i(e_k) = \sum_{k=1}^{\infty} |a_i(k)|$$

and $\varphi_i(b_k) < \infty$ since $b_k \in l^\infty$. The infinite sum above is valid by continuity of φ_i. Now given an arbitrary $x(k) \in l^\infty$ we show that $\sum_{k=1}^{\infty} a_i(k) x(k) \to 0$ as $i \to \infty$. Note

$$\varphi_i(x) = \sum_{k=1}^{\infty} \varphi_i(e_k)\, x(k) = \sum_{k=1}^{\infty} a_i(k) x(k)$$

and since $\varphi_i \to 0$ weak*-ly, $\varphi_i(x) \to 0$ as $i \to \infty$, and we have proved the claim, since $a_i(k)$ converges to 0 weakly in l^1, and thus in l^1 norm as well.

Note that continuity of π was needed to show that $\varphi_i \in (l^\infty)'$

Interpolation Theory

Besides Open Mapping Theorem, there are other methods to verify that some linear maps are bounded. In Harmonic Analysis for instance, a basic tool is **interpolation**. Suppose $A: L^p \to L^q$. In many cases, it is easy to show that A is bounded for specific choices for p and q. Given these, interpolation tells us that A is bounded for choices of p and q "in-between" the choices that we can prove easily. Here we prove one such interpolation theorem. It is not the most useful, but can be applied to various operators.

Theorem 94. (Riesz-Throin) *Let $p_0 \geq 1$ and $p_1, q_0, q_1 \leq \infty$, $0 < \theta < 1$. Suppose that T is a linear operator from $L^{p_0}(X) + L^{p_1}(X) \to L^{q_0}(X) + L^{q_1}(X)$, and that*

$$|Tu|_{L^{q_0}} \leq A_0 |u|_{L^{p_0}} \text{ and } |Tu|_{L^{q_1}} \leq A_1 |u|_{L^{p_1}}$$

i.e. T is bounded from $L^{p_0} \to L^{q_0}$ and from $L^{p_1} \to L^{q_1}$. Then we have that

$$|Tu|_{L^{q_\theta}} \leq A_0^{1-\theta} A_1^{\theta} \, |u|_{L^{p_\theta}}$$

where

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \text{ and } \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$$

i.e. that T is bounded from $L^{p_\theta} \to L^{q_\theta}$.

Before the proof, we consider some applications of this interpolation theorem.

Example 95. (Convolution) Consider the operator

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y) g(y) \, dy = \int_{\mathbb{R}^n} f(y) g(x-y) \, dy$$

Then if $1 \leq p, q, r \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$, then

$$|f * g|_{L^r} \leq |f|_{L^p} + |g|_{L^q}$$

Note there are two easy cases. For instance, if $r = \infty$, then $\frac{1}{p} + \frac{1}{q} = 1$ and

$$|f * g|_\infty \leq \int |f(x-y)| \, |g(y)| \, dy \leq \|f\|_p \|g\|_q$$

by Hölder's inequality. Also, if $r = 1$, then $p = q = 1$ and

$$|f * g|_1 \leq \int \int |f(x-y)| \, |g(y)| \, dy \, dx \leq \int \|f\|_1 \, |g(y)| \, dy \leq \|f\|_1 \, \|g\|_1$$

by Fubini. Then, using interpolation,

1. Using the operator $T_f(g) = f * g$, with $f \in L^1$, we know that $|T_f \, g|_1 \leq |f|_1 \, |g|_1$ and that $|T_f \, g|_\infty \leq |f|_1 |g|_\infty$, we can show that $|T_f \, g|_p \leq |f|_1 \, |g|_p$. In the theorem, take $p_0 = q_0 = 1$ and $p_1 = q_1 = \infty$, then for θ,

   $$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} = 1 - \theta$$

   so that $p_\theta = \frac{1}{1-\theta}$. The same holds for $q_\theta = p_\theta$. Riesz-Thorin tells us that

   $$|T_f \, g|_{p_\theta} \leq |f|_1^{1-\theta} |f|_1^{\theta} \, |g|_{p_\theta} = |f|_1 \, |g|_{p_\theta}$$

   Thus we have $|f * g|_p \leq |f|_1 \, |g|_p$.

2. For the full range of p, q, we apply interpolation again. Using $T_g f = g * f$ this time, note that $|T_g \, f|_q \leq |g|_q \, |f|_1$ from above and that $|T_g \, f|_\infty \leq |g|_q \, |f|_{q'}$. Then we apply the theorem with $p_0 = 1$, $q_0 = q$ and $p_1 = q'$, $q_1 = \infty$. Then for θ,

   $$\frac{1}{p_\theta} = 1 - \theta + \frac{\theta}{q} = 1 - \frac{\theta}{q'}$$

and

$$\frac{1}{q_\theta} = \frac{1-\theta}{q'}$$

Riesz-Thorin tells that

$$|T_g f|_{q_\theta} \leq |g|_q^{1-\theta} |g|_q^\theta |f|_{p_\theta} = |g|_q |f|_{p_\theta}$$

so that $|g * f|_{q_\theta} \leq |g|_q |f|_{p_\theta}$ where

$$1 + \frac{1}{q_\theta} = 1 + \frac{1-\theta}{q'} = \frac{1}{q} + \frac{1}{p_\theta}$$

as desired.

Now we turn to the proof of Riesz-Thorin.

Proof. First, note $L^{p_\theta} \subset L^{p_0} + L^{p_1}$ (cutoff the function u at some level λ. Then $u\mathbf{1}_{\{|u|<\lambda\}} \in L^{p_1}$ and $u\mathbf{1}_{\{|u|>\lambda\}} \in L^{p_0}$) Given $u \in L^{p_\theta}$, we want to estimate $|Tu|_{q_\theta}$. Note by duality, and density of simple functions,

$$|Tu|_{q_\theta} = \sup_{\substack{|v|_{q_\theta'} \leq 1 \\ v \,\text{simple}}} \int_Y (Tu)\, v\, dy$$

It also suffices to bound $|Tu|$ for simple functions u, because otherwise we can take simple u_n converging to u in L^{p_θ}, and $u_n = u_n' + u_n''$ with u_n' converging to some u' in L^{p_0} and u_n'' converging to some u'' in L^{p_1}. This implies by boundedness on L^{p_0} and L^{p_1} that $Tu_n' \to Tu'$ in L^{q_0} and $Tu_n'' \to Tu''$ in L^{q_1}. Then we can find a subsequence $u_{n'}$ of u_n for which $Tu_n \to Tu$ almost everywhere, and then we apply Fatou's lemma so that

$$\int (Tu)\, v\, dy \leq \liminf_{n\to\infty} \int (Tu_n)\, v\, dy$$

so that the bound for $|T\,u|$ for nonsimple functions can be obtained with bounds for simple functions. Now define

$$F(z) = \int_Y T\left(|u|^{p_\theta/p(z)} \frac{u}{|u|}\right) |v|^{q_\theta'/q(z)'} \frac{v}{|v|}\, dy$$

where $|v|_{L^{q_\theta'}} \leq 1$, and $\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}$ and $\frac{1}{q(z)} = \frac{1-z}{q_0} + \frac{z}{q_1}$, $\frac{1}{q(z)'} = 1 - \frac{1}{q(z)}$. Note that noting that $z = 0$, $|u|^{p_\theta/p_0} \in L^{p_0}$ and $|v|^{q_\theta'/q_0'} \in L^{q_0}$, and likewise for p_1 and q_1. For u, v simple, F is holomorphic, since for instance if $v = \sum_k v_k \mathbf{1}_{B_k}$, then

$$|v|^{q_\theta'/q(z)'} = \sum_k |v_k|^{q_\theta'/q(z)'} \mathbf{1}_{B_k}(y) \frac{|v_k|}{v_k}$$

where z only appears in the exponent terms. Also, for $0 \leq \operatorname{Re} z \leq 1$, $|F(z)| \leq C$ bounded since the sums are finite. Then since $|F(0+it)| \leq A_0$ (apply $T: L^{p_0} - L^{q_0}$) and $|F(1+it)| \leq A_1$ (apply $T: L^{p_1} - L^{q_1}$), we apply the Hadamard's three lines theorem on the region $\{z \in \mathbb{C}: 0 \leq \operatorname{Re} z \leq 1\}$ to get that

$$|F(\theta)| \leq A_0^{1-\theta} A_1^\theta$$

but $F(\theta) = \int_Y T(u)\, v\, dy$ so we are done after taking the supremum over v with norm 1.

□

Locally Convex Spaces

There are two motivations for considering locally convex spaces.

- Earlier we introduced weak convergence, and a natural question to ask is, "Under which topology is convergence the same as weak convergence?" The result will not be a normed space.

- We know that $C(K, \mathbb{C})$ for K compact is normed with the L^∞ norm. What about $C(\Omega, \mathbb{C})$ for $\Omega \subset \mathbb{R}^n$ open? We already have notions of convergence, for instance, uniform on every compact subset. The corresponding topology will not be a normed space here either.

Weak Topology

Let X be a Banach space. Recall weak convergence, where $x_n \to x$ weakly if $l(x_n) \to l(x)$ for all $l \in X'$. We now investigate which topology corresponds to this notion of convergence.

In general, let $f_\alpha: X \to Y_\alpha$ $\alpha \in \Lambda$ be a family of functions mapping from a set X to a topological space Y_α. Consider

$$S = \{ f_\alpha^{-1}(V_\alpha) : V_\alpha \subset Y_\alpha \text{ is open in } Y_\alpha \}$$

preimages of open sets. Then let τ_S be the smallest topology containing S. We call τ_S the topology generated by f_α.

Remark 96. A few notes:

- τ_S is the weakest topology on X for which all f_α are continuous maps from $(X, \tau_S) \to Y_\alpha$. This is by definition of τ_S. f_α is continuous means $f_\alpha^{-1}(V)$ is open for V open in Y_α, and we already set up S so that $f_\alpha^{-1}(V) \in S$.

- If $x_j, x \in X$, then $x_j \to x$ in τ_S if and only if $f_\alpha(x_j) \to f_\alpha(x)$ for all α.

  Proof. Note that if $x_j \to x$ in τ_S, then $f_\alpha(x_j) \to f_\alpha(x)$ for all α since f_α are continuous from $(X, \tau_S) \to Y_\alpha$. Conversely, suppose that $f_\alpha(x_j) \to f_\alpha(x)$ for all α. We want to show that for any open neighborhood $U(x)$ (containing x) in (X, τ_S), that $x_j \in U$ for all sufficiently large j. Since τ_S is generated by S, there exist α_i for which

  $$x \in f_{\alpha_1}^{-1}(V_{\alpha_1}) \cap \cdots \cap f_{\alpha_n}^{-1}(V_{\alpha_n}) \subset U(x)$$

  This implies that $f_{\alpha_i}(x) \in V_{\alpha_1}$. But then since $f_{\alpha_i}(x_j) \to f_{\alpha_i}(x)$ for $\alpha_1, \ldots, \alpha_n$, we see that for sufficiently large j, $f_{\alpha_i}(x_j) \in V_{\alpha_1}$ as well. This implies that

  $$x_j \in f_{\alpha_1}^{-1}(V_{\alpha_1}) \cap \cdots \cap f_{\alpha_n}^{-1}(V_{\alpha_n}) \subset U(x)$$

  for all sufficiently large j. Thus $x_j \to x$ in (X, τ_S). □

This tells us how to generate the topology consistent with the weak topology.

Weak and Weak* Toplogies:

For the Banach space X, and its dual $X' = \{l: X \to \mathbb{C} \text{ bounded}\}$, define

$$\sigma(X, X') := \text{topology on } X \text{ generated by } X'$$

(above, X is the same, and X' is the family of functions from $X \to \mathbb{C}$). By the remarks above, convergence in this topology is the same as weak convergence.

Likewise, for X', if we consider the family $\varphi_x: X' \to \mathbb{C}$ defined by $\varphi_x(l) = l(x)$, we denote

$$\sigma(X', X) := \text{ topology on } X' \text{ generated by } (\varphi_x)_{x \in X}$$

And $l_j \to l$ in this topology if and only if $l_j(x) = \varphi_x(l_j) \to \varphi_x(l) = l(x)$ for all $x \in X$. This is simply the weak* topology.

Remark 97. The weak topology does not have a countable neighborhood base (i.e. is not A_1) if $\dim X = \infty$. Recall for (X, τ) topological space, having a countable neighborhood base means that for every $x \in X$, we can find a countable collection of open neighborhoods $V_j(x)$ around x for which for which every neighborhood $U(x)$ contains some $V_j(x)$.

The reason this is significant is that if (X, τ) has a countable neighborhood base, then a mapping $f: X \to Y$ is continuous if and only if for every $x_j \to x$ in (X, τ) we have that $f(x_j) \to f(x)$.

Note: If f is continuous, then $f(x_j) \to f(x)$ for every sequence $x_j \to x$ already. Take any neighborhood around $f(x)$, the preimage is open, containing x, and thus x_j is inside the preimage for j sufficiently large, and thus $f(x_j)$ is in the neighborhood for all j sufficiently large.

The issue is the converse. If we know $f(x_j) \to f(x)$ for all $x_j \to x$, it is not necessarily true that f is continuous. But if (X, τ) has a countable neighborhood base, it is true. It is a consequence of the fact that for such spaces, x is in the closure of A if and only if there exists a sequence in A converging to x.... and etc. See texts on topology / wikipedia.

This issue is resolved in general by the introduction of "nets", which generalize sequences.

Proof. (of Remark) Let $X/\mathbb{R}$ be a real Banach space, let $\dim X = \infty$, and consider the weak topology $(X, \sigma(X, X'))$. Suppose $(V_i)_{i=1}^{\infty}$ is a base for 0, then for all j, by definition of the weak topology there exists $l_{i,1}, \cdots, l_{i,m_i} \in X'$ for which $\bigcap_{j=1}^{m_i} l_{i,j}^{-1}(-1,1) \subset V_i$ (by scaling we can use the preimage of the same set $(-1,1)$). Then, given any $l \in X'$, since $\{V_i\}$ is a countable base for 0, for some V_i we have that

$$\bigcap_{j=1}^{m_i} l_{i,j}^{-1}(-1,1) \subset V_i \subset l^{-1}(-1,1)$$

This implies that for all $\lambda > 0$,

$$\bigcap_{j=1}^{m_i} l_{i,j}^{-1}(-\lambda,\lambda) \subset l^{-1}(-\lambda,\lambda)$$

and hence taking $\lambda \to 0$,

$$\bigcap_{j=1}^{m_i} N(l_{i,j}) \subset N(l)$$

There are now two claims to be made.

1. The null space containment above implies that l is a linear combination of $l_{i,j}$.

2. If X is a Banach space and $\dim X = \infty$, X cannot be expressed as the span of a countable basis S (i.e. not every x can be written as a finite linear combination of elements of S).

Given the first claim, every l is in the span of $\{l_{i,1}, ..., l_{i,m}\}$ for some i, and thus

$$X' = \operatorname{span}\{l_{i,j}: 1 \leq i \leq \infty, 1 \leq j, \leq m_i\}$$

But given the second claim, this is a contradiction.

Proof of First Claim: Suppose $l_1, ..., l_n$ and l are in X' and $\bigcap_{i=1}^{n} N(l_i) \subset N(l)$. Then construct a mapping $A: X \to \mathbb{R}^n$ by $A(x) = (l_1(x), ..., l_n(x))$. Then we wish to show that there exists a linear functional φ: $\mathbb{R}^n \to \mathbb{R}$ for which $\varphi \circ A = l$, in which case

$$l(x) = \varphi(l_1(x), ..., l_n(x)) = \sum_{i=1}^{n} \varphi(e_i) l_i(x)$$

so that l is a linear combination of the l_i. Note that $A\,x = 0$ implies that $l(x) = 0$, so $N(A) \subset N(l)$. This means that it is well defined to define $\varphi(y) = l(A^{-1}(y))$. Note that $A^{-1}(y) = x + N(A)$ for some x. Then

$$l(A^{-1}(y)) = l(x + N(A)) = l(x)$$

so that this is well defined. Clearly it is linear, so we are done.

Proof of Second Claim: We prove the contrapositive. Suppose that there exists a countable basis e_i for which every $x \in X$ can be written as a finite linear combination of e_i. Note that $K_n = \operatorname{span}\{e_1, ..., e_n\}$ is a sequence of closed subspaces of X (since finite dimensional) and whose union is X. By Baire's Category Theorem (Corollary 63), one of the K_n has nonempty interior, $B_r \subset K_n$. But this implies that $X \subset K_n$ by linearity, and since K_n is finite dimensional, X is finite dimensional. □

Thus, the weak topology $(X, \sigma(X, X'))$, without a countable neighborhood base does not necessarily have an easy characterization of continuous functions from $X \to Y$ in terms of sequences.

Nets

First, a **directed set** is a set I and a relation $\leq$ on I such that:

- $\alpha \leq \alpha$
- (Transitive) $\alpha \leq \beta$, $\beta \leq \gamma \Longrightarrow \alpha \leq \gamma$
- For all $\alpha, \beta \in I$, there exists $\gamma \in I$ such that $\alpha \leq \gamma$ and $\beta \leq \gamma$.

Example 98. $(\mathbb{N}, \leq)$ is a directed set.

Example 99. For (X, τ), consider $x_0 \in X$, then if we take

$$I = \{U(x_0): U \text{ neighborhood of } x_0\}$$

then $(I, \supset)$ is a directed set, i.e. $U \leq V$ if $U \supset V$. In fact, this is what we will be using to construct "nets."

Example 100. If $(I, \leq_I)$ and $(J, \leq_J)$ are directed sets, then so is $(I \times J, \leq)$ where

$$(\alpha_1, \alpha_2) \leq (\beta_1, \beta_2) \Longleftrightarrow \alpha_1 \leq \beta_1 \text{ and } \alpha_2 \leq \beta_2$$

A **net** in X is a mapping from a directed set $(I, \leq) \to X$, where we will be using the notation $x_\alpha \in X$ with $\alpha \in I$. Note that a special case is sequences, where we use $x_n \in X$ with the directed set $(\mathbb{N}, \leq)$.

Net convergence: In (X, τ), let $(\Lambda, \leq)$ be a directed set, and consider the net $(x_\alpha)_{\alpha\in\Lambda} \in X$. Then we say that $x_\alpha \to x$ iff for all $U(x)$, there exists α_0 such that $x_\alpha \in U(x)$ for $\alpha \geq \alpha_0$.

Proposition 101. *Let $f: (X, \tau_X) \to (Y, \tau_Y)$ be a mapping, and let $x_0 \in X$. Then f is continuous at x_0 if and only if $f(x_\alpha) \to f(x_0)$ for all convergent nets $x_\alpha \to x_0$.*

Proof. First suppose f is continuous at x_0. Then let $V = V(f(x_0))$ some neighborhood of $f(x_0)$. Then $f^{-1}(V)$ contains some neighborhood $U(x_0)$ of x_0. Now suppose $x_\alpha \to x_0$. This means that there is some α_0 for which $x_\alpha \in U(x_0)$ for all $\alpha \geq \alpha_0$. But since $U(x_0) \subset f^{-1}(V)$ this implies that $f(x_\alpha) \in V$ for all $\alpha \geq \alpha_0$ and therefore $f(x_\alpha) \to f(x_0)$.

Conversely, suppose that f is not continuous at x_0. Then there exists $V = V(f(x_0))$ for which given any neighborhood $U(x_0)$, there exists $x_U \in U(x_0)$ with $f(x_U) \notin V$. Note that using the directed set $(\{U(x_0)\}, \supset)$, x_U is a net with $x_U \to x_0$, since given any $U'(x_0)$, $x_U \in U'$ for all $U \subset U'$ $(U \geq U')$. However, given any $U'(x_0)$, $f(x_{U'}) \notin V$, so $f(x_U) \not\to f(x_0)$. $\square$

Two facts written in notes, not sure if relevant:

1. *X is Hausdorff if and only if the (net) limit is only unique, i.e. if $x_\alpha \to x$ and $x_\alpha \to x'$ then $x = x'$.*

2. *(There exists?) $f_\alpha: X \to Y_\alpha$ for $\alpha \in \Lambda$ for which*

$$x_\beta \to x \Longleftrightarrow f_\alpha(x_\beta) \to f_\alpha(x) \quad \textit{for all } \alpha$$

Subnets. Given $(x_\alpha)_{\alpha\in I}$ a net, then $(y_\beta)_{\beta\in I}$ is a subnet of x_α if there exists $\varphi: J \to I$ for which

1. $y_\beta = x_{\varphi(\beta)}$

2. For all $\alpha_0 \in I$, there exists $\beta_0 \in J$ such that $\beta \geq \beta_0 \Longrightarrow \varphi(\beta) \geq \alpha_0$.

Properties:

1. In (X, τ),

$$\overline{A} = \{x \in X, \text{ exists } x_\alpha \in A \text{ s.t. } x_\alpha \to x\}$$

2. If $\{x_\alpha, \alpha \in I\}$ has x as a limit point, then there exists a subnet $(y_\beta)_{\beta\in J}$ for which $y_\beta \to x$.

3. X is compact if and only if every net in X has a convergent subnet.

Reference: See Reed-Simon, Chapter 2.

In summary, the main idea is that in a topological space, we can describe continuity of functions with preimages of open sets, but *also* as the preservation of net convergence.

Topological Vector Spaces

A **topological vector space** is a vector space X with a topology (X, τ) such that addition and scalar multiplication are continuous maps. In other words,

1. Addition, $X \times X \to X$ with $(x, y) \mapsto x + y$ is continuous

2. Scalar multiplication, $\mathbb{K}\times X\to X$ with $(\lambda,x)\mapsto\lambda x$ is continuous.

Banach spaces $(X,|\cdot|)$ with topology generated by the norm is a topological vector space (addition and scalar multiplication are continuous with respect to the norm).

Example 102. Now that we have characterized continuity with net convergence, we can show easily that $(X,\sigma(X,X'))$ is a topological vector space. For instance, if we take a net on $X\times X$, with $(x_\alpha,y_\alpha)\to(x,y)$ weakly, which means that $x_\alpha\to x$ and $y_\alpha\to y$ (product topology), then by the definition of weak topology, for all $l\in X'$, $l(x_\alpha)\to l(x)$ and $l(y_\alpha)\to l(y)$. But by linearity, this implies that $l(x_\alpha+y_\alpha)\to l(x+y)$, and thus $x_\alpha+y_\alpha\to x+y$ weakly, and this implies that addition is continuous with respect to the weak topology, since net convergence is preserved. The same computation holds for scalar multiplication.

Proposition 103. *Let X be a Hausdorff topological vector space, and let $\dim X=n<\infty$. Then X is isomorphic to $\mathbb{R}^n$*

Proof. Let $u_1,\ldots,u_n$ be a basis for X. Define $\varphi:\mathbb{R}^n\to X$ by $(t_1,\ldots,t_n)\mapsto\sum_{i=1}^n t_i u_i$. Note φ is continuous because addition and scalar multiplication are continuous. We want to show that φ is a homeomorphism (preserves topology). Let $\xi\in\mathbb{R}^n$, and let $S=\{|\xi|=1\}$. Then S is compact, and $\varphi(X)\subset X$ is also compact, therefore closed. Since $0\notin\varphi(S)$, we can find $V=V(0)$ such that $V\cap\varphi(X)=\varnothing$, i.e. $V\subset X\backslash\varphi(S)$. Now for every $0\le t\le 1$, $t\,V\subset V$, and the claim is that this implies $\varphi(B_1)\supset V$.

This is because given $x\in V$, $x=\varphi(\xi)$ for some ξ, and since for $0\le t\le 1$, $\varphi(t\,\xi)=t\,x\notin\varphi(S)$ (since $t\,x\in V$ and V does not intersect with $\varphi(S)$), we have that $t\,\xi\notin S$ for any $0\le t\le 1$ so that $|\xi|<1$, so $x\in\varphi(B_1)$. ☐

Locally Convex Spaces

A **locally convex space** is a topological vector space (X,τ) where every neighborhood $U(0)$ contains an open convex neighborhood $V(0)$ of 0.

For instance, Banach spaces are locally convex since $B_r=\{|x|<r\}$ are convex, and the weak topology $(X,\sigma(X,X'))$ is also a locally convex space since every neighborhood contains a set of the form

$$l_1^{-1}(-\varepsilon_1,\varepsilon_1)\cap\cdots l_k^{-1}(-\varepsilon_k,\varepsilon_k)$$

which is convex and open.

Example 104. Another example is furnished by a family of seminorms. Let X be a real vector space, and let p_α for $\alpha\in\Lambda$ be a family of seminorms (norms, but without the property that $|x|=0\Rightarrow x=0$). We define the topology

$$\tau=\left\{U\colon\ \text{for all } y\in U,\ \text{there exists } \alpha_i,r_i,1\le i\le n \text{ s.t. } \bigcap_{i=1}^n B_{r_i}^{p_{\alpha_i}}(y)\subset U\right\}$$

where $B_r^p(y)=\{x\colon p(x-y)<r\}$ (seminorm ball)

(As a special case, the weak topology is generated by the family of seminorms $p_l(x)=|l(x)|$ for $l\in X'$)

Then (X,τ) is locally convex. Note that Remark 96 and the same proof as in Example 102 show that (X,τ) is a topological vector space. By definition of τ, every open neighborhood contains a finite intersection of seminorm balls, which are convex. Also, Remark 96 shows that all p_α are continuous with respect to (X,τ).

Proposition 105. *Every locally convex space is generated by a family of seminorms.*

Proof. Let X be a locally convex space. Let $V \subset X$ be a convex neighborhood of 0 with $-V = V$ (symmetric). Recall the Minkowski functional (gauge) from Remark 4:

$$p_V(x) = \inf_{\substack{\lambda > 0 \\ x \in \lambda V}} \lambda$$

This is a seminorm: Let $a = p_V(x), b = p_V(y)$. Then $x \in (a + \varepsilon)V$ and $y \in (b + \varepsilon)V$for all ε. Then

$$x + y \in (a + b + 2\varepsilon)V$$

for all ε, and thus $p_V(x + y) \leq a + b + 2\varepsilon$ for all ε. Thus p_V is subadditive, and hence a seminorm (since it is nonnegative and $p_V(0) = 0$). Now we claim that $(X, \tau) = (X, (p_V)_{V \text{ convex, symmetric nbd}})$, i.e. that the Minkowski functionals generate the topology of X.

$(\subset)$ Let $U = U(0)$ be an open neighborhood in (X, τ). Then there exists a convex set V with $V \subset U$, and furthermore we can make V symmetric by replacing V with $V' = (-V) \cap V$. Then

$$B_1^{p_{V'}}(0) \subset V \subset U$$

Note $x \in B_1^{p_{V'}}$ means that $x \in \alpha V \subset V$ some $\alpha < 1$. Thus U is in the topology generated by the seminorms p_V.

$(\supset)$ Let W be a neighborhood of 0 in $(X, (p_V))$. Then

$$\bigcap_{i=1}^{n} \{p_{v_i}(x) < \varepsilon_i\} \subset W$$

where the LHS is a finite intersection of open sets in (X, τ), and hence is open in (X, τ). $\square$

Example 106. Let $\Omega \subset \mathbb{R}^n$ be open, and let $X = C(\Omega, \mathbb{R})$. Recall that X is not normable. However, we can generate the topology of uniform convergence on compact sets by a family of seminorms. Let $K \subset \Omega$, and define

$$p_K(\varphi) = \sup_{x \in K} |\varphi(x)|$$

Then $(p_K)_{K \subset \Omega}$ generate a topology on $C(\Omega)$, where

$$\varphi_\alpha \xrightarrow{C(\Omega)} \varphi \Longleftrightarrow \varphi_\alpha|_K \longrightarrow \varphi|_K \text{ uniformly}$$

In fact, $C(\Omega)$ has a countable neighborhood base, and can be made into a metric space (Take a countable sequence K_n which increases to Ω, with $K_j \subset \text{int}(K_{j+1})$, and use $d(\varphi, \psi) = \sum_n \frac{p_{K_n}(\varphi - \psi)}{1 + p_{K_n}(\varphi - \psi)} 2^{-n}$. It is an exercise to show that the two topologies are equivalent, i.e. $(X, (p_K)_j) = (X, d)$)

Frechet Spaces

A **Frechet Space** is a locally convex Hausdorff space which has a countable neighborhood base, and every Cauchy sequence has a limit. In fact, under these conditions, such a space is metrizable under a translation invariant metric, and under this metric the space is a complete metric space. (See Reed-Simon, Ch 3-4).

For instance, $(C(\Omega, \mathbb{R}), d)$ in the previous example is a Frechet Space.

Example 107. Consider the Schwarz class,

$$\mathcal{S}(\mathbb{R}^n) = \left\{ \varphi \in C^\infty(\mathbb{R}^n) \text{ s.t. } \sup_{\substack{x \in \mathbb{R}^n \\ |\alpha| \le m}} (1 + |x|)^m |\partial^\alpha \varphi(x)| < \infty \right\}$$

For instance, $\varphi(x) = e^{-|x|^2}$ is in $\mathcal{S}(\mathbb{R}^n)$

Taking $p_m(\varphi) = \sup_{\substack{x \in \mathbb{R}^n \\ |\alpha| \le m}} (1 + |x|)^m |\partial^\alpha \varphi(x)|$ which are seminorms, $(\mathcal{S}(\mathbb{R}^n), (p_m)_{m=0}^\infty)$ is a Frechet Space.

Fourier Transform and Tempered Distributions.

For $\varphi \in \mathcal{S}$ define

$$\hat{\varphi}(\xi) = \int_{\mathbb{R}^n} \varphi(x)\, e^{-2\pi i x \cdot \xi}\, d\,\xi$$

Recall $\mathcal{F}: \varphi \mapsto \hat{\varphi}$ is an isomorphism from $\mathcal{S} \to \mathcal{S}$. The space of continuous linear functionals on $\mathcal{S}$ is called the space of tempered distributions $\mathcal{S}'$. By duality we can find $\mathcal{F}'$ so that the diagram commutes:

$$\begin{array}{ccc} \mathcal{S}' & \xrightarrow{\mathcal{F}'} & \mathcal{S}' \\ \uparrow & & \uparrow \\ \mathcal{S} & \xrightarrow{\mathcal{F}} & \mathcal{S} \end{array}$$

where the map $\mathcal{S} \to \mathcal{S}'$ is the usual mapping taking $\varphi \in \mathcal{S}$ to the functional $(\psi \mapsto \int \varphi\, \psi) \in \mathcal{S}'$.

i.e. that $[\mathcal{F}'(l)](\varphi) = l(\mathcal{F}\varphi)$ for $l \in \mathcal{S}'$. So we have a “Fourier transform” $\mathcal{F}'$ on the space of tempered distributions which agrees with the usual Fourier transform on L^1, L^2 and on measures $\mu(\mathbb{R}^n)$. For instance,

$$\begin{array}{ccc} L^1(\mathbb{R}^n) & \longrightarrow & \mathcal{S}' \\ \wedge \big\downarrow & & \big\downarrow \mathcal{F} \\ C_0(\mathbb{R}^n) & \longrightarrow & \mathcal{S}' \end{array}$$

commutes. The same situation occurs with L^2 and the space of measures.

We can define “derivatives” for distributions in a similar manner. Recall

$$\widehat{\partial_j \varphi} = i\, \xi_j\, \hat{\varphi}(\xi)$$

call the natural map $J: \mathcal{S} \to \mathcal{S}'$ where $J\varphi(\psi) = \int \varphi\, \psi$. Note that J is injective, and integration by parts gives

$$\langle J(\partial_k \varphi), \psi \rangle = (-1)^k \langle J(\varphi), \partial_k \psi \rangle$$

(here $\langle \cdot, \cdot \rangle$ denotes functional evaluation, not inner product). Then for $T \in \mathcal{S}'$,

$$\langle \partial_k T, \psi \rangle = (-1)^k \langle T, \partial_k \psi \rangle$$

and in particular, for $T = J(\varphi)$ we have $\partial_k J(\varphi) = J(\partial_k \varphi)$.

Further Reference: Rudin, *Functional Analysis*, Part II.

Example 108. *Distribution Spaces*. Let $\Omega \subset \mathbb{R}^n$, and let

$$\mathcal{D}(\Omega) = C_c^\infty(\Omega, \mathbb{C})$$

Let $K\subset\Omega$ be compact, and define

$$\mathcal{D}_K(\Omega)=\left\{\varphi\in\mathcal{D}(\Omega)\colon \varphi|_{\Omega\setminus K}=0\right\}$$

$$p_m(\varphi)=\sup_{\substack{x\in K\\|\alpha|\leq m}}|\partial^\alpha\varphi(x)|$$

and so $(\mathcal{D}_K(\Omega),\ (p_m)_{m=0}^\infty)$ is a Frechet Space. However, the construction of the topology for $\mathcal{D}(\Omega)$ is strange. Let

$$S=\left\{\mathrm{co}\left(\bigcup_{K\subset\Omega\text{ compact}}U_K(0)\right), U_K(0)\text{ is some neighborhood of 0 in }\mathcal{D}_K\right\}$$

(convex hull)

Define

$$\tau=\{U\subset\mathcal{D}(\Omega),\ \text{for all }\varphi\in U,\ \text{exists}\,V\in S\text{ s.t. }\varphi+V\subset U\}$$

Then

- $(\mathcal{D}(\Omega),\tau)$ is a locally convex space
- S is a neighborhood base of 0
- $\mathcal{D}_K(\Omega)\xrightarrow{i_K}\mathcal{D}(\Omega)$, and τ is the largest locally convex topology on $\mathcal{D}(\Omega)$ such that all i_K are continuous.
- If $T\colon\mathcal{D}(\Omega)\to Y$, where Y is locally convex, is a linear map, then T is continuous if and only if $T|_{\mathcal{D}_K}$ is continuous for all K.
- $\mathcal{D}(\Omega)$ does not have a countable neighborhood base. $(\mathcal{D}(\Omega),\tau)$ is called the "inductive topology".

Why is it so complicated? It turns out that

$$(\mathcal{D}(\Omega),(p_{m,K})_{m\in\mathrm{N},K\subset\Omega})\subset(C^\infty(\Omega),(p_{m,K})_{m\in\mathrm{N},K\subset\Omega})$$

where $p_{m,K}(\varphi)=\sup_{|\alpha|\leq m,x\in K}|\partial^\alpha\varphi(x)|$ is not complete.

($\mathcal{D}'$ is the space of distributions).

Separation Theorem

Returning to general locally convex spaces, recall that in the earlier treatment of Hahn-Banach we showed the existence of linear functions bounded above by a postiive homogeneous and subadditive function $p(x)$. Using this we have shown the existence of nontrivial continuous linear functionals on *normed* spaces and also separation of convex sets with continuous linear functionals. Now in the weaker setting of locally convex spaces, we can again show this with an appropriate choice for $p(x)$. The key to show continuity of the linear functional in a locally convex space is to show continuity at 0, that if $x\to 0$ (**Note! We are using net convergence here**), then $l(x)\to 0$ as well.

Proposition 109. *If X is a locally convex space and X_0 is a subspace with $l_0\in X_0'$, then there exists $l\in X'$ such that $l|_{X_0}=l_0$.*

Proof. Note there exists U open and convex such that $0 \in U$ and

$$U \cap X_0 \subset \{|l_0| < 1\}$$

Then

$$|l_0(x)| \leq p_U(x)$$

for $x \in X_0$. This is because

$$\begin{aligned} x \in \lambda U &\implies \left|l_0\left(\frac{x}{\lambda}\right)\right| \leq 1 \\ &\implies |l_0(x)| \leq \lambda \\ &\implies |l_0(x)| \leq p_U(x) \end{aligned}$$

Then by Hahn-Banach there exists $l: X \to \mathbb{R}$ for which $l(x) \leq p_U(x)$ and $l|_{X_0} = l_0$. Then

$$|l(x)| \leq \max\{p_U(x), p_{-U}(x)\}$$

and therefore if $x \to 0$ then $l(x) \to 0$, so that $l \in X'$. $\square$

Example 110. Let X be locally convex Hausdorff space, and $x_0 \in X$ nonzero. Then there exists $l \in X'$ such that $l(x_0) = 1$ by extension of the finite dimensional (and hence continuous) map $t x_0 \mapsto t$ from $\mathbb{K} X_0 \to \mathbb{K}$. By linearity we also have $l(t x_0) = t$.

Proposition 111. *Let X be a locally convex space over $\mathbb{R}$, and $A, B \subset X$ be convex subsets with $\operatorname{int} A \neq 0$ and $A \cap B = \varnothing$. Then there exists $l \in X'$ such that $l \neq 0$ and $l(x) \leq l(y)$ for all $x \in A$, $y \in B$. (For complex vector spaces, obtain $\operatorname{Re} l \to l$).*

Proof. This is exactly Theorem 3, using the previous proposition for the extension instead. $\square$

Double Polar Theorem

Let $\langle \cdot, \cdot \rangle$ be a bilinear function $X \times Y \to \mathbb{R}$, not neceesarily an inner product. Then we say we have a **dual pair** of spaces X, Y. We say that (X, Y) is **separate** (or nondegenerate) if

$$\langle x, y \rangle = 0 \text{ for all } y \in Y \Longrightarrow x = 0$$

and the same for the other coordinate.

Weak Toplogies. Given the bilinear function $\langle \cdot, \cdot \rangle$, for all y we can define a linear functional l_y on X by $l_y(x) = \langle x, y \rangle$. Then $(l_y)_{y \in Y}$ generates a locally convex space $(X, \sigma(X, Y)) = (X, (l_y)_{y \in Y})$. The same is true for $(Y, \sigma(Y, X))$.

Note: Given a net x_α, note that

$$x_\alpha \to x \text{ in } \sigma(X, Y) \Longleftrightarrow \langle x_\alpha, y \rangle \to \langle x, y \rangle \text{ for all } y \in Y$$

Therefore if (X, Y) is separate, then the spaces $(X, \sigma(X, Y))$ and $(Y, \sigma(X, Y))$ are Hausdorff (nets have a unique limit).

Example 112. As special cases, we already know (X, X') as a separate dual pair with bilinear function

$$\langle x, l \rangle = l(x)$$

then $(X,\, \sigma(X,\, X'))$ is the weak topology on X and $(X',\, \sigma(X',\, X))$ is the weak* topology on X'. Also, (X', X'') is a separate dual pair with $(X', \sigma(X', X''))$ the weak topology on X'.

Dual Topologies.

Proposition 113. *Let (X, Y) be a separate dual pair. Then $(X, \sigma(X, Y))' = Y$*

Proof. Note if $y \in Y$, then $l_y(x) = \langle x,\, y\rangle$ defines a linear functional in $(X, \sigma(X, Y))'$, so we have that $Y \subset (X, \sigma(X, Y))'$. Now suppose $l \in (X, \sigma(X, Y))'$. Then we know that

$$\{|l| < 1\} \supset \bigcap_{j=1}^{m} \{|l_{y_j}| < \varepsilon_j\}$$

since the topology is generated by $(l_y)_{y \in Y}$. The same proof as in Remark 97 shows that

$$N(l) \supset \bigcap_{j=1}^{m} N(l_{y_j})$$

and that

$$l = c_1 l_{y_1} + \ldots + c_m l_{y_m} = l_{c_1 y_1 + \ldots + c_m y_m}$$

i.e. that l can be identified with $c_1 y_1 + \ldots + c_m y_m \in Y$. $\square$

Definition of Polar. Let (X, Y) be a dual pair with $\langle \cdot\,, \cdot \rangle: X \times Y \to \mathbb{R}$. Let $A \subset X$ and $B \subset Y$. Then we define

$$\begin{aligned} A^\circ &= \{y \in Y: \langle x,\, y\rangle \leq 1 \text{ for all } x \in A\} \\ {}^\circ B &= \{x \in X: \langle x,\, y\rangle \leq 1 \text{ for all } y \in B\} \end{aligned}$$

called the **polar** of A and the **polar** of B. There is nothing special about which side the $\circ$ is, since we can swap X, Y (and the bilinear function) to reverse them. This is just a notational convenience.

Note, we will use $\operatorname{int} A$ to denote the interior of A, and A° to denote the polar of A

Example 114. For X Banach, and X, X' the dual pair with $\langle x,\, l\rangle = l(x)$ as before, we note that denoting B_1^X by the unit ball in X, $B_1^X = \{x \in X: |x| \leq 1\}$, we have that

$$\begin{aligned} (B_1^X)^\circ &= B_1^{X'} \\ {}^\circ(D_1^{X'}) &= B_1^X \end{aligned}$$

by duality principle (Proposition 48), since $|x| = \sup_{l \in X', |l| \leq 1} |l(x)|$, and $|l| = \sup_{x \in X, |x| \leq 1} |l(x)|$.

Also, if $Y \subset X$ a subspace, then

$$Y^\circ = \{l \in X': l(y) = 0 \text{ for all } y \in Y\}$$

because if $l(t\,y) \leq 1$ for all t, then $l(y) = 0$.

Observations:

- $A^\circ \subset Y$ is convex, and $0 \in A^\circ$.

- A° is closed in $\sigma(Y, X)$. This follows since A° is the intersection of closed sets $\{y: \langle x, y\rangle \leq 1\}$ (inverse image of a closed set for a continuous linear functional)

Theorem 115. (Double Polar Theorem) *Given (X, Y) a separate dual pair with $\langle \cdot, \cdot \rangle$. For $A \subset X$,*

$$^\circ(A^\circ) = \overline{\text{co}(A \cup \{0\})}$$

where co *denotes convex hull and the closure is with respect to $\sigma(X, Y)$.*

Proof. ($\supset$) First note that $A \subset {}^\circ(A^\circ)$, since if $a \in A$, then for all $A \in A^\circ$ by definition $\langle a, y\rangle \leq 1$ and hence $a \in {}^\circ(A^\circ)$. Also, $0 \in {}^\circ(A^\circ)$ trivially. Furthermore, $^\circ(A^\circ)$ is convex and closed in $\sigma(Y, X)$ and therefore taking the convex hull and then the closure shows that

$$\overline{\text{co}(A \cup \{0\})} \subset {}^\circ(A^\circ)$$

($\subset$) Towards a contradiction, suppose we have $x_0 \in {}^\circ(A^\circ)$ and $x_0 \notin \overline{\text{co}(A \cup \{0\})}$. Then there exists U open in $\sigma(X, Y)$ such that $x_0 + U \cap \overline{\text{co}(A \cup \{0\})} = \varnothing$. Since we are in a locally convex space, U can be chosen to be convex. This implies that there exists $l \in (X, \sigma(X, Y))'$ nonzero such that

$$l(z) \leq l(x_0 + w)$$

where $z \in \overline{\text{co}(A \cup \{0\})}$ and $w \in U$ by the separation theorem. Since $l \neq 0$ there exists some $w \in U$ for which $l(w) > 0$ (pick arbitrary $w \in U$ and scale appropriately, may need U to be balanced as well, which is fine). Then

$$l(z) \leq l(x_0) - l(w) < l(x_0)$$

for all $z \in \overline{\text{co}(A \cup \{0\})}$, and in particular, using $z = 0$ shows that $l(x_0) > 0$.

Thus there exists $c > 0$ for which $l(z) \leq c < l(x_0)$ and by scaling, $c^{-1} l(z) \leq 1 < c^{-1} l(x_0)$. Since $(X, \sigma(X, Y))' = Y$ (identification), we have that $c^{-1} l = l_y$ for some $y \in Y$, so that

$$\langle z, y\rangle \leq 1 < \langle x_0, y\rangle$$

again for all $z \in \overline{\text{co}(A \cup \{0\})}$, which implies on one hand that $y \in A^\circ$, but on the other hand that $x \notin {}^\circ(A^\circ)$, which contradicts our assumption on x.

□

Remark 116. Note that the theorem also works for $(^\circ B)^\circ$ by symmetry (swapping X, Y for instance)

Corollary 117. *If X is a uniformly convex Banach space, then X is reflexive.*

Proof. Recall that uniformly convex means that for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x| = |y| = 1$ and $\left|\frac{x + y}{2}\right| > 1 - \delta$ then $|x - y| < \varepsilon$.

Let $\kappa: X \to X''$ be the natural identification map given by $\kappa x(l) = l(x)$. We wish to show that κ is onto. Let $B_1^X = \{x \in X: |x| \leq 1\}$, and likewise for $B_1^{X'}$ and $B_1^{X''}$. Consider the dual pair (X', X'') with $\langle l, \varphi\rangle = \varphi(l)$. Note $\kappa(B_1^X) \subset X''$. Now

$$^\circ(\kappa(B_1^X)) = \{l \in X': \kappa x(l) \leq 1 \text{ for all } |x| \leq 1\} = \{l \in X': |l(x)| \leq 1 \text{ for all } |x| \leq 1\} = B_1^{X'}$$

and that by the Double Polar Theorem (Theorem 115),

$$\overline{\kappa(B_1^X)} = ({}^\circ(\kappa(B_1^X)))^\circ = (B_1^{X'})^\circ = B_1^{X''}$$

where the closure above is in $\sigma(X'', X')$. Now if $\varphi \in X''$, with $|\varphi|_{X''} = 1$, we need to find x such that $\kappa x = \varphi$. Since $\varphi \in B_1^{X''} = \overline{\kappa(B_1^X)}$, we can find $x_\alpha \in B_1^X$ (net) with $\kappa x_\alpha \to \varphi$ in $\sigma(X'', X')$ (the weak* sense). We want to show that x_α is a Cauchy net, from which we conclude that $x_\alpha \to x$ for some x (Exercise, but should follow from simply extracting a Cauchy sequence from the Cauchy net).

Note $\liminf_\alpha |\kappa x_\alpha| \geq |\varphi| = 1$ by lower semicontinuity under weak* convergence, and thus $|x_\alpha| \to 1$ since $x_\alpha \in B_1^X$. Replace x_α by $\frac{x_\alpha}{|x_\alpha|}$, and since $|x_\alpha| \to 1$ we still have $\kappa \frac{x_\alpha}{|x_\alpha|} \to \varphi$ in the weak* sense. Thus now we have x_α unit vectors, and note

$$\kappa\left(\frac{x_\alpha + x_\beta}{2}\right) \to \varphi$$

in the weak* sense as well. Then

$$1 \leq \liminf_{\alpha,\beta} \left|\frac{x_\alpha + x_\beta}{2}\right| \leq \limsup_{\alpha,\beta} \left|\frac{x_\alpha + x_\beta}{2}\right| \leq 1$$

so that $\left|\frac{x_\alpha + x_\beta}{2}\right| \to 1$. By uniform convexity, this implies that $|x_\alpha - x_\beta| \to 1$, so that x_α is a Cauchy net. Therefore $x_\alpha \to x$ for some x and $\kappa x = \varphi$.

□

Krein Milman Theorem

Theorem 118. *Let X be a locally convex Hausdorff space, and $K \subset X$ a convex, compact subset. Then*

1. *The extremal set of K is nonempty.*

2. *$K = \overline{\mathrm{co}(\mathrm{ext}(K))}$ (i.e. K is the closure of the convex hull of the extremal points of K)*

Before we prove the theorem, we make a few remarks and applications.

Remark 119. If $K \subset \mathbb{R}^n$, then there is a simple result by Carathéodory which states that every compact subset K in $\mathbb{R}^n$ has extreme points, and every point of K can be written as a convex combination of $n+1$ extremal points.

The proof is by induction on the dimension n. For $n=1$ the statement is trivial, as K is just an interval. Suppose the result is true for all subspaces of dimension less than n. There are two cases for K.

- If K has empty interior in $\mathbb{R}^n$, then K is contained in a lower dimensional affine subspace, and (translating to a linear subspace) by induction we are done.

- Otherwise, first we claim that every boundary point of K can be represented as a convex combination of n extreme points of K. Given $x \in \partial K$, we can find a linear functional l for which $l(x) \leq l(y)$ for all $y \in K$. Consider $K_x = \{y \in K \colon l(y) = l(x)\}$, which is a hyperplane intersected with the boundary of K. K_x is an extreme subset of K, as all other points z satisfy $l(z) < l(x)$.

  Since this is lower dimensional, by induction every point in K_x can be written as a convex combination of n extreme points of K_x. Note that the extreme points of K_x are also extreme points of K. Thus any $x \in \partial K$ can be written as a convex combination of n extreme points of K.

To finish the proof, now consider an arbitrary point $y \in K$. If y is an extreme point already, there is nothing to prove. Otherwise, pick an arbitrary extreme point x_1 of K, and form the line $\{z: z = t\, x_1 + (1-t)y,\, t \in \mathbb{R}\}$. This line necessarily intersects ∂K at some x_2. By above, x_2 is a convex combination of n extreme points. Since y is contained in the line segment connecting x_1 and x_2, then y is a convex combination of x_1 and x_2, and thus y is a convex combination of x_1 and n other extreme points, and we are done.

Where to get Compactness

Proposition 120. (Banach-Alaoglu) *If X is Banach, then $B_1^{X'}$, the unit ball in the dual is compact in the weak* topology $\sigma(X', X)$.*

Proof. Consider the map $\varphi: B_1^{X'} \to \prod_{\substack{|x| \le 1 \\ x \in X}} [-1, 1]$ defined by $\varphi(l) = (l(x))_{x \in X}$. The infinite product is compact by Tychonoff's Theorem (the topology is the topology of pointwise convergence). Then

1. φ is a homeomorphism.

$$\begin{aligned} l_\alpha \xrightarrow{\sigma(X',X)} l &\iff l_\alpha(x) \to l(x) \text{ for all } x \in X \\ &\iff l_\alpha(x) \to l(x) \text{ for all } |x| \le 1, x \in X \\ &\iff \varphi(l_\alpha) \to \varphi(l) \text{ pointwise} \end{aligned}$$

2. $\varphi(B_1^{X'})$ is closed.

   Suppose that $\varphi(l_\alpha) \to f$, so that $l_\alpha(x) \to f(x)$ for all $|x| \le 1$. Then for $x \neq 0$, let $f(x) = f\left(\frac{x}{|x|}\right)|x|$ (extension). Thus

$$l_\alpha(x) = |x|\, l_\alpha\left(\frac{x}{|x|}\right) \to |x|\, f\left(\frac{x}{|x|}\right) = f(x)$$

   so that $l_\alpha(x) \to f(x)$ for all x. By linearity of limits, f is linear, and also $|f(x)| \le 1$ for $|x| \le 1$ since f is in the product of $[-1, 1]$. Thus $f \in B_1^{X'}$, and $f = \varphi(f)$.

This shows that $\varphi(B_1^{X'})$ is a closed subset of a compact space, and hence is compact, and since φ is a homeomorphism, $B_1^{X'}$ is also compact. $\square$

Application of Krein Milman

Theorem 121. *$L^1(0, 1)$ is not the dual of any space Y.*

Proof. Suppose towards a contradiction that $L^1 \cong Y'$. Then consider $B_1^{L^1}$, by the previous Proposition 120 $B_1^{L^1}$ is compact in the weak* $\sigma(L^1, Y)$ topology. However, we note that $\operatorname{ext}(B_1^{L^1}) = \varnothing$. Note given any $f \in B_1^{L^1}$, we can split $f = f\mathbf{1}_{[0,b]} + f\mathbf{1}_{[b,1]}$ where b satisfies

$$\int_0^b |f| = \int_b^1 |f| = \frac{|f|_{L^1}}{2}$$

(since $\int_0^x |f| dx$ is continuous in x, we can find such a b). Then we have that

$$f = \frac{1}{2}\left(2f\mathbf{1}_{[0,b]}\right) + \frac{1}{2}\left(2f\mathbf{1}_{[b,1]}\right)$$

where $|2f\mathbf{1}_{[0,b]}|_{L^1} = |2f\mathbf{1}_{[b,1]}|_{L^1} = |f|_{L^1}$, i.e. both parts are in $B_1^{L^1}$ and so f is not extremal.

Now by Krein-Milman, (Theorem 118), we have that $B_1^{L^1}$ being compact in $\sigma(L^1, Y)$ is the closure of the convex hull of its extreme points, but since the set of extreme points is empty, this is a contradiction. Thus L^1 cannot be the dual of any space.

(A technical detail is that we need to identify L^1 homeomorphically to Y', but this is not an issue since we assumed $L^1 \cong Y'$) □

Proof of Krein Milman

Proof. (of Theorem 118) Let $\mathcal{E} = \{L\colon L \text{ is a closed, extremal subset of } K\}$. Noting that $K \in \mathcal{E}$, we know that $\mathcal{E}$ is nonempty. Now we have an ordering on $\mathcal{E}$ defined by

$$L_1 \leq L_2 \Longleftrightarrow L_1 \supset L_2$$

Then every linearly ordered subset of $\mathcal{E}$ has an upper bound, since the intersection is an upper bound, which is in $\mathcal{E}$ since the intersection of nonempty compact sets is also nonempty and compact. Also, the intersection of extremal subsets is also extremal. Thus by Zorn's Lemma, there exists a maximal element K_0 in $\mathcal{E}$. The claim is that $K_0 = \{x\}$, i.e. consists of a single point. This implies that $x \in \operatorname{ext}(K)$ so that $\operatorname{ext}(K) \neq \varnothing$.

Suppose that K_0 consists of multiple points, and let $x_1, x_2 \in K_0$ with $x_1 \neq x_2$. Then there exists a linear functional $l \in X'$ for which $l(x_1) < l(x_2)$. Then let $K_1 = \{x \in K_0\colon l(x) = \max_{K_0} l\}$ which is also an extremal subset of K_0 and hence an extremal subset of K. Since $l(x_1) < l(x_2)$, $x_1 \notin K_1$ and hence $K_1 > K_0$, contradicting the maximality of K_0.

Now we show that $K = \overline{\operatorname{co}(\operatorname{ext}(K))}$. Suppose not, then let $x_0 \in K \setminus \overline{\operatorname{co}(\operatorname{ext}(K))}$. Then there exists $l \in X'$ such that $l(x) \leq c < l(x_0)$ for all $x \in \overline{\operatorname{co}(\operatorname{ext}(K))}$. Then define $K_0 = \{x \in K\colon l(x) = \max_K l\}$. This is an extremal subset of K, which is nonempty by continuity of l. Now since K_0 is nonempty and convex, take $y \in \operatorname{ext}(K_0)$ (exists by previous part). Then y is also an extreme point of K (extreme point of an extreme subset is still an extreme point of the original set). However, this is a contradiction since $y \notin \operatorname{ext}(K)$ by hyperplane separation: points in $\operatorname{ext}(K)$ satisfy $l(y) \leq c < l(x_0) \leq \max_K l$, and hence cannot be in K_0.

□

Spectrum of Bounded Linear Operators

Recall from Linear Algebra that for a matrix A,

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\Longleftrightarrow \lambda I - A \text{ not invertible} \\ &\Longleftrightarrow \det(\lambda I - A) = 0 \\ &\Longleftrightarrow N(\lambda I - A) \neq \{0\} \end{aligned}$$

In particular, A has exactly n eigenvalues, and we have

- In general, have the Jordan Form
- If $A^* = A$ (self-adjoint), then $A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*$ for U unitary and λ_j real.
- If $A^*A = A\,A^*$ (normal), then $A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*$ for U unitary and λ_j complex.

We hope to have similar results for operators between Banach spaces under some conditions.

Let $X/\mathbb{C}$ be a Banach space, and let $A \in \mathcal{L}(X)$ (bounded operator from $X \to X$). We define

$$\begin{aligned}\sigma(A) &:= \{\lambda \in \mathbb{C}: \lambda I - A \text{ is not invertible}\} \\ &= \{\lambda \in \mathbb{C}: N(\lambda I - A) \neq \{0\}\} \bigcup \{x \in \mathbb{C}: R(\lambda I - A) \neq X\}\end{aligned}$$

$\sigma(A)$ is the spectrum, and it consists of two parts, the first part $\{\lambda, \lambda I - A \neq \{0\}\}$ are the eigenvalues or point spectrum, which we denote $\sigma_p(A)$, and the second part will remain unnamed.

The complement, $\rho(A) = \mathbb{C} \backslash \sigma(A)$ is called the resolvent, and

$$\begin{aligned}\rho(A) &= \{\lambda \in \mathbb{C}: \lambda I - A \text{ is invertible}\} \\ &= \{\lambda \in \mathbb{C}: N(\lambda I - A) = \{0\}, R(\lambda I - A) = X\}\end{aligned}$$

Example 122. Let $X = C([0,1], \mathbb{C})$ with the supremum norm $|f|_\infty = \max_{0 \leq x \leq 1} |f(x)|$, and consider the multiplication operator $A(f) = x f$. This is a bounded linear operator. Consider the equation

$$(\lambda I - A) f = (\lambda - x) f$$

First we note that if $(\lambda I - A) f = 0$, then $f(x) = 0$ for $x \neq \lambda$, and by continuity $f \equiv 0$. Thus $N(\lambda I - A) = \{0\}$ for all λ, and thus A has no point spectrum (eigenvalues).

Note that if $\lambda \notin [0,1]$, then it is easy to solve $(\lambda - x) f = g$ by $f = \frac{g}{\lambda - x}$ (which is continuous now that $\lambda \notin [0,1]$). Then for $\lambda \notin [0,1]$, $\lambda I - A$ is one-to-one and onto, and hence bijective. Thus $\mathbb{C} \backslash [0,1] \subset \rho(A)$

On the other hand, if $\lambda \in [0,1]$, note that $(\lambda I - A) X \subset \{f \in C[0,1], f(\lambda) = 0\}$ which is a strict subspace of X, and thus $[0,1] \subset \sigma(A)$.

This implies that $\sigma(A) = [0,1]$ and $\rho(A) = \mathbb{C} \backslash [0,1]$.

Example 123. Let $H = L^2[0,1]$, and study the multiplication operator again $A f = x f$.

If $(\lambda I - A) f = 0$, then $(\lambda - x) f(x) = 0$ for all x. Then $(\lambda - x) f(x) = 0$ except on a set of measure zero, and since $\lambda - x$ is only zero at $x = \lambda$, this implies that $f(x)$ must be zero except on a set of measure zero, so $f = 0$ (in L^2).

From the same reasoning as above, we have that if $\lambda \notin [0,1]$ then $(\lambda I - A) \frac{g}{\lambda - x} = g$ so that $\mathbb{C} \backslash [0,1] \subset \rho(A)$. Now suppose $\lambda \in [0,1]$. Then we show that $\lambda I - A$ is not onto. More specifically, if we choose $g = 1$ (constant function), then there does not exist $f \in L^2[0,1]$ such that

$$(\lambda I - A) f = 1$$

Otherwise, we must have $f = \frac{1}{\lambda - x}$ which is not in L^2. Thus $\sigma(A) = [0,1]$ and $\rho(A) = \mathbb{C} \backslash [0,1]$ and A has no point spectrum since the null space is trivial for all λ.

Holomorphic Functions from $\mathbb{C} \to X$

Before we turn to facts about $\sigma(A)$, we develop a tool that will be useful in many of the proofs to follow.

We will be using results in complex analysis. Recall that a function $f: \mathbb{C} \to \mathbb{C}$ is holomorphic at z_0 if

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

exists (the limit is in $\mathbb{C}$). Likewise, we can use the same definition for functions from $\mathbb{C}\to X$. A function $f:\mathbb{C}\to X$ is (strongly) holomorphic at z_0 if

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

where the limit is in X.

A function $f:\mathbb{C}\to X$ is weakly holomorphic at z_0 if for every linear functional $l\in X'$, $l\circ f$ is holomorphic as a function from $\mathbb{C}\to\mathbb{C}$. We can also see this as the same definition as above but replacing the limit in X with the weak limit. With this view, it is clear that strongly holomorphic functions are weakly holomorphic.

For Banach spaces, it turns out that if f is weakly holomorphic on some open domain Ω, then it is also strongly holomorphic on Ω. (Also true for Frechet spaces)

Proposition 124. *If f is weakly holomorphic in an open domain Ω, then it is strongly holomorphic in Ω.*

Proof. See Lax, Section 11.4 □

Many results in complex analysis will carry through.

- (Cauchy's Theorem) The Cauchy-Goursat Theorem says that if f is holomorphic in some domain Ω, then

  $$\int_\gamma f(z)\,dz=0$$

  for γ a closed curve in Ω. There is a proof that uses solely the differentiability of f and first proving the Theorem for rectangles and repeatedly cutting the rectangle into fourths.

  The Cauchy Integral Formula

  $$f(z)=\frac{1}{2\pi i}\int_\gamma\frac{f(\xi)}{\xi-z}d\xi$$

  can be deduced from this when applying the Cauchy-Goursat Theorem to $\frac{f(\xi)-f(z)}{\xi-z}$.

  Thus, Liouville's Theorem, Residue Theorem, Power Series, Laurent Series all hold. Below we present different approaches for the proofs.

- (Liouville's Theorem) Suppose that $f(z):\mathbb{C}\to X$ is entire (holomorphic on $\mathbb{C}$) and bounded. Then the same is true for $l\circ f$, for any l, and thus f must also be constant.

- (Power Series) A useful tool is that if we can express $f(z)$ has a power series $\sum_{k=0}^{\infty} c_k\, z^k$ where $c_k\in X$, then f is holomorphic within the radius of convergence (recall in a Banach space X, any absolutely summable series is summable).

- (Laurent Series) We can use power series to transfer to Laurent series which are valid in an annulus $r_1<|z|<r_2$, $f(z)=\sum_{k=-\infty}^{\infty} c_k\, z^k$ with $c_k\in X$. Furthermore, we can still talk about residues, $\mathrm{Res}_{z=0}f(z)=c_1$, and

  $$\frac{1}{2\pi i}\int_\gamma f(z)\,dz=\mathrm{Res}_{z=0}f(z)$$

This is justified by term by term integration, and using the usual complex analysis identity (Residue Theorem)

$$\frac{1}{2\pi i}\int_\gamma c_k z^k\,dz = \frac{c_k}{2\pi i}\int_\gamma z^k\,dz = \begin{cases} c_k & k=-1 \\ 0 & \text{else} \end{cases}$$

Facts about $\sigma(A)$

Now we turn to basic facts about the spectrum $\sigma(A)$.

Proposition 125. *Let $A \in \mathcal{L}(X)$. Then*

1. *$\sigma(A)$ is compact*

2. *$\sigma(A) \neq \varnothing$*

3. *The spectral radius, $r_A := \sup_{z\in\sigma(A)} |z| = \max_{z\in\sigma(A)} |z|$ (max since $\sigma(A)$ is compact) satisfies*

$$r_A = \lim_{n\to\infty} |A^n|^{1/n}$$

We will be using the following fact:

Proposition 126. *If $|A| < 1$, then $I - A$ is invertible, and*

$$(I-A)^{-1} = \sum_{j=0}^{\infty} A^j$$

The same formula holds if $|A|^m < 1$ for some m as well.

Proof. Since $|A^j| \leq |A|^j$ and $|A| < 1$, the RHS series is absolutely summable, and hence converges to some element. Note that a computation involving telescoping series shows that

$$(I-A)\sum_{j=0}^{n-1} A^j = \left(\sum_{j=0}^{n-1} A^j\right)(I-A) = I - A^n$$

and as $m \to \infty$, we have that $I - A^n \to I$. (note $|A^n| \leq |A|^n \to 0$). Thus

$$\sum_{j=0}^{\infty} A^j = (I-A)^{-1}$$

Now to show that the same formula holds when $|A|^m < 1$ for some m, we simply show that the RHS series is absolutely summable, and then the same computation carries through (replacing n by multiples of m). Write $\theta = |A^m|^{1/m}$. Then for any j note we can write $j = km + l$ so that

$$\begin{aligned} |A^j| &= |A^{km+l}| \\ &\leq |A^m|^k\,|A^l| \\ &= \theta^{mk+l}\cdot\frac{|A^l|}{\theta^l} \\ &\leq C\theta^j \end{aligned}$$

where $C = \max_{0 \le l \le m} \frac{|A^l|}{\theta^l}$.

□

Now we return to the facts about the spectrum $\sigma(A)$.

Proof. (of Proposition 125)

(1) Compactness of $\sigma(A)$**:** First we show that $\sigma(A)$ is compact, and since it is a subset of $\mathbb{C}$, it suffices to show that $\sigma(A)$ is closed and bounded. Note that for $|z| > |A|$, we have that $z\,I - A = z(I - z^{-1}A)$ is invertible, so $z \in \rho(A)$ and (taking complements) $\sigma(A) \subset \overline{B_{|A|}}$. Thus $\sigma(A)$ is bounded.

Now we show that $\rho(A)$ is open, which shows $\sigma(A)$ is closed. Let $z_0 \in \rho(A)$. Then

$$z\,I - A = (z_0 I - A)\big[\,I + (z - z_0)(z_0 I - A)^{-1}\big]$$

where for $|z - z_0|$ sufficiently small (smaller than $|(z_0 I - A)^{-1}|$), $I + (z - z_0)(z_0 I - A)^{-1}$ is invertible, and hence $z\,I - A$ is invertible for all $|z - z_0|$ sufficiently small. Thus we can find $B_\varepsilon(z_0) \subset \rho(A)$ and hence $\rho(A)$ is open

(2) $\sigma(A)$ **nonempty:** To show that $\sigma(A)$ is nonempty, we will use the fact that $z\,I - A$ is holomorphic on $\rho(A)$ (as a function from $\mathbb{C} \to X$, see previous section). This is because for any $z_0 \in \rho(A)$, by the representation above we can expand $(z\,I - A)^{-1}$ in a power series about z_0 for $|z - z_0| < \varepsilon$, ε sufficiently small:

$$(z\,I - A)^{-1} = \left[\sum_{j=0}^{\infty} (-1)^j (z_0 I - A)^{-j} (z - z_0)^j\right] (z_0 I - A)^{-1}$$

Since we can express $(z\,I - A)^{-1}$ as a power series around any $z_0 \in \rho(A)$, $(z\,I - A)^{-1}$ is holomorphic on $\rho(A)$. Now, suppose towards a contradiction that $\sigma(A) = \varnothing$, or $\rho(A) = \mathbb{C}$, so that $(z\,I - A)^{-1}$ is entire.

The claim is that $(z\,I - A)^{-1}$ is also bounded. First note that for any linear operator $B \in \mathcal{L}(X)$, if

$$\inf_{|x|=1} |B\,x| = \rho$$

then

$$|B^{-1}|_X = \frac{1}{\rho} \sup_{|x|=\rho} |B^{-1}\,x| \le \frac{1}{\rho}$$

In particular, for $|z| > 2|A|_X$, we have that $\inf_{|x|=1} |(z\,I - A)x| \ge |z| - |A|_X$ so that

$$|(z\,I - A)^{-1}|_X \le \frac{1}{|z| - |A|_X} \le \frac{1}{|z|}$$

For $|z| \le 2|A|_X$, we note that $z \mapsto |(z\,I - A)^{-1}|_X$ is continuous, and hence is bounded on the compact set $|z| \le 2\,|A|_X$. Thus $(z\,I - A)^{-1}$ is bounded. By Liouville's Theorem (generalized), $(z\,I - A)^{-1}$ reduces to a constant, and since by the above estimate $|(z\,I - A)^{-1}|_X \to 0$ as $z \to \infty$, $(z\,I - A)^{-1}$ must be identically 0, which is a contradiction (0 is not invertible). Thus it must be the case that $\rho(A) \neq \mathbb{C}$, so $\sigma(A)$ is nonempty.

(3) Spectral Radius: We will show this in three steps.

Step 1. First we show that the limit exists, and in particular

$$\lim_{n \to \infty} |A^n|^{1/n} = \inf_{m \ge 0} |A^m|^{1/m}$$

Now fixing m, let $n = k m + l$ for $0 \le l \le m - 1$. Then

$$\begin{aligned} |A^n| &\le |A^{km+l}| \\ &\le |A^m|^k |A^l| \\ &= |A^m|^{n/m} \cdot \frac{|A^l|}{|A^m|^{l/m}} \\ &\le C\,|A^m|^{n/m} \end{aligned}$$

where $C = \max_{0 \le l \le m-1} \frac{|A^l|}{|A^m|^{l/m}}$. Taking n-th roots gives $|A^n|^{1/n} \le C^{1/n} |A^m|^{1/m}$, and thus taking limsup in n and inf over m we have

$$\limsup_{n \to \infty} |A^n|^{1/n} \le \inf_{m \ge 0} |A^m|^{1/m} \le \liminf_{n \to \infty} |A^n|^{1/n}$$

and the right inequality follows from $\liminf_{n \to \infty} |A^n|^{1/n} = \sup_n \inf_{m > n} |A^m|^{1/m}$. Furthermore, since liminf is bounded above by limsup, we have that

$$\liminf_{n \to \infty} |A^n|^{1/n} = \limsup_{n \to \infty} |A^n|^{1/n} = \inf_{m \ge 0} |A^m|^{1/m}$$

Step 2. Now we show that $r_A \le \lim_{n \to \infty} |A^n|^{1/n}$. It suffices to show that if $|z| > \lim_{n \to \infty} |A^n|^{1/n}$ then $z \in \rho(A)$. If $|z| > \lim_{n \to \infty} |A^n|^{1/n}$ then for some m we have that $|z| > |A^m|^{1/m}$ for some m. We can rewrite this as $\left|\left(\frac{A}{z}\right)^m\right|^{1/m} < 1$. Now by Proposition 126, this implies $I - \frac{A}{z}$ is invertible, in which case $z I - A$ is invertible so that $z \in \rho(A)$

Step 3. Finally we show that $r_A \ge \lim_{n \to \infty} |A^n|^{1/n}$. For $|z| > |A|$ we have the Laurent series expansion

$$\left(I - \frac{A}{z}\right)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{z^k}$$

or

$$(z I - A)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}}$$

and if we integrate over the curve $\{|z| = R\} \subset \rho(A)$

$$\frac{1}{2\pi i} \int_{|z| = R} z^k (z I - A)^{-1}\, dz = A^k$$

(see previous section) Then we have the bound

$$|A^k| \le \frac{1}{2\pi} \int_0^{2\pi} C(R)\, R^{k+1}\, d\theta = C(R)\, R^{k+1}$$

with $|(z I - A)^{-1}| \le \sum_{k=0}^{\infty} \frac{1}{|z|} \cdot \left|\frac{A}{z}\right|^k = C(R)$. Taking k-th roots and taking the limit as $k \to \infty$ shows that

$$\lim_{k \to \infty} |A^k|^{1/k} \le R$$

and since $\{|z| = R\} \subset \rho(A)$ for all $R > r_A$, letting $R \to r_A$ we conclude that

$$\lim_{k \to \infty} |A^k|^{1/k} \le r_A$$

$\square$

Example 127. Note that if A is nilpotent, i.e. $A^m = 0$, then $r_A = 0$. This is easily seen in the case where $X = \mathbb{R}^m$ and

$$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

where all the eigenvalues are 0.

Functional Calculus

As in finite dimensional linear algebra, we can define $A^2 = A \cdot A$ and if $p(z) = \sum_{k=0}^{n} c_k z^k$ then

$$p(A) = \sum_{k=0}^{n} c_k A^k$$

Furthermore, if f is entire with power series $f(z) = \sum_{k=0}^{\infty} c_k z^k$, we can define

$$f(A) = \sum_{k=0}^{\infty} c_k A^k$$

For instance, the exponential of an operator is

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

We can even go further and define $f(A)$ for f holomorphic from $\Omega \to \mathbb{C}$ where Ω contains $\sigma(A)$ via the integral formula

$$f(A) := \frac{1}{2\pi i} \int_\gamma f(\xi) (\xi I - A)^{-1} d\xi$$

where $n(\gamma, z) = 1$ for all $z \in \sigma(A)$ and $n(\gamma, z) = 0$ for all $z \notin \Omega$. (This integral can be defined as a limit of Riemann sums). This corresponds to the Residue Formula in complex analysis, and will allow us to define for instance $\log(A)$ in a region containing $\sigma(A)$ that is simply connected and does not contain 0.

Note in the case that f is entire, then we can use a power series to see that this definition matches with the previous. Choose γ to be $|z| = R$ with $R > |A|$, so that

$$\begin{aligned} \frac{1}{2\pi i} \int_\gamma f(z) (zI - A)^{-1} &= \frac{1}{2\pi i} \int_\gamma f(z) \left(\sum_{k=0}^{\infty} z^{-k-1} A^k \right) dz \\ &= \sum_{k=0}^{\infty} A^k \left(\frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z^{k+1}} dz \right) \\ &= \sum_{k=0}^{\infty} c_k A^k \\ &= f(A) \end{aligned}$$

Now in finite dimensions, we know that if A is $n \times n$ matrix, and $\lambda_1, \ldots, \lambda_n$ are its eigenvalues, then for f entire, we know that $f(\lambda_1), \ldots, f(\lambda_n)$ are the eigenvalues of $f(A)$. This follows from computing $f(A)$ of the Jordan form of A. The diagonal terms of $f(A)$ are precisely $f(\lambda_k)$.

A natural question is to ask in the general setting, whether $\sigma(f(A)) = f(\sigma(A))$.

Proposition 128. *Let $A \in \mathcal{L}(X)$, and let f, g be holomorphic functions from $\Omega \to \mathbb{C}$ with Ω containing $\sigma(A)$. Then*

1. $f(A)\, g(A) = (f g)(A)$

2. $\sigma(f(A)) = f(\sigma(A))$

3. $g(f(A)) = (g \circ f)(A)$, *where g is holomorphic in a region containing $\sigma(f(A))$ instead.*

Proof. First we show that $(1) \Longrightarrow (2)$. Assume that $f(A)\, g(A) = (f g)(A)$. First we show that $f(\sigma(A)) \subset \sigma(f(A))$ Let $z_0 \in \sigma(A)$. We want to show that $f(z_0) \in \sigma(f(A))$. Define $g(z) = \frac{f(z) - f(z_0)}{z - z_0}$ which is holomorphic for the same region as f (the singularity at $z = z_0$ is removable). Now

$$f(z) - f(z_0) = (z - z_0)\, g(z)$$

and applying (1) we have

$$f(A) - f(z_0)\, I = (A - z_0\, I)\, g(A) = g(A)\, (A - z_0\, I)$$

and since $A - z_0 I$ is not invertible, $f(A) - f(z_0) I$ is not invertible either: If $N(A - z_0\ I) \neq \{0\}$, then $N(g(A)(A - z_0\ I)) \neq \{0\}$, and if $R(A - z_0 I) \neq X$, then $R((A - z_0 I) g(A)) \neq X$. This implies that $z_0 \in \sigma(f(A))$ so that $(\sigma(A)) \subset \sigma(f(A))$.

To show that $\sigma(f(A)) \subset f(\sigma(A))$, suppose that $w \notin f(\sigma(A))$. We will show that $w\, I - f(A)$ is invertible so that $w \notin \sigma(f(A))$. If we choose $g(z) = \frac{1}{w - f(z)}$, then g is holomorphic in a neighborhood of $\sigma(A)$. Then since $(w - f(z)) g(z) = 1$, we apply (1) to get

$$(w\, I - f(A)) g(A) = g(A) (w\, I - f(A)) = I$$

so that $w\, I - f(A)$ is invertible (left and right inverse).

For a proof of (1) and (3), see Lax, Theorem 5 in Section 17.2. □

Fredholm Theory

Note: This is a rearrangement of material covered in class.

In finite dimensions, if $A: X \to X$ is a linear map, then we have the following two properties:

1. A is injective if and only if A is surjective.

2. We have that

$$\begin{aligned} A x = b \quad \Longrightarrow \quad & \langle A x, y \rangle = \langle b, y \rangle \text{ for all } y \in X \\ & \left\langle x, A^T y \right\rangle = \langle b, y \rangle \text{ for all } y \in X \end{aligned}$$

   so that if x is a solution to $A x = b$, then a necessary condition is that $\langle b, y \rangle = 0$ for all y such that $A^T y = 0$, i.e. that $b \in (N_{A^T})^{\perp}$. This turns out to be a sufficient condition also. In other words,

$$R_A = (N_{A^T})^{\perp}$$

This is almost immediate when examining the matrices, particularly the fact the operation Ax can be viewed as taking the inner product of the rows of A with x:

$$R_A = \mathrm{col}(A) = \mathrm{row}(A^T) = (N_{A^T})^{\perp}$$

(A technical note: to use complex inner product we use the adjoint A^* instead of the transpose)

However, these properties may not hold in infinite dimensions:

Example 129. Let $l^2 = \{(a_n)_{n=1}^{\infty}: \sum_{n=1}^{\infty} |a_n|^2 < \infty\}$, and define $A: l^2 \to l^2$ by

$$A(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$$

(the right shift operator). Note that A is injective, but not surjective. This is an example of a property in finite dimensions that does not hold in infinite dimensions (the equivalence of injectivity and surjectivity).

The second property still holds: The range looks like $(0, a_1, a_2, \ldots)$. A^T is precisely the left shift, since

$$\begin{aligned} \langle A(a_1, a_2, \ldots), (b_1, b_2, \ldots)\rangle &= \langle (0, a_1, a_2, \ldots), (b_1, b_2, \ldots)\rangle \\ &= \langle (a_1, a_2, \ldots), (b_2, b_3, \ldots)\rangle \\ &= \big\langle (a_1, a_2, \ldots), A^T(b_1, b_2, \ldots)\big\rangle \end{aligned}$$

Then the null space of A^T consists of elements of the form $(r, 0, 0, \ldots)$, and the orthogonal complement is precisely the range of A, with elements of the form $(0, a_1, a_2, \ldots)$.

Example 130. Consider $X = L^2([0,1]) = \left\{f: \int_0^1 |f(x)|^2\, dx < \infty\right\}$, and the multiplication operator M defined by

$$Mf(x) = x f(x)$$

Note that M is a self-adjoint operator:

$$\int Mf(x)\, \overline{g(x)}\, dx = \int x f(x)\, \overline{g(x)}\, dx = \int f(x)\, \overline{Mg(x)}\, dx$$

Note that $N_M = \{0\}$, so that $(N_M)^{\perp} = X$, but $R_M \neq X = (N_M)^{\perp}$ since $1 \notin R_M$.

Later, however, we will show that it is the case that $\overline{R_M} = \ker(N_M)^{\perp}$.

A natural question to consider, then, is for which operators A does the properties (1) and (2) above remain true? Fredholm operators enjoy the same properties, and we will see them shortly.

Index and Pseudoinverse

Note: For what follows, X, Y are generic vector spaces.

Let $K: X \to Y$ be a linear operator, with X, Y vector spaces over $\mathbb{R}$. If $\mathrm{ran}(K)$ is finite dimensional, then we say that K is of **finite rank**. Note the following properties:

- If $X \xrightarrow{A} Y \xrightarrow{B} Z$, then $B \circ A$ is of finite rank if either B or A is of finite rank.
- If $K_1, K_2: X \to Y$ are of finite rank, then $K_1 + K_2$ is of finite rank.

Pseudoinverse. If $A: X \rightarrow Y$ is a linear map, and there exists $B: Y \rightarrow X$ such that

$$\begin{aligned} AB &= I + K \\ BA &= I + L \end{aligned}$$

where K, L are of finite rank, then we say that A has a pseudoinverse, and B is called the pseudoinverse of A. Note that if the dimensions are finite, every map is pseudoinvertible.

Why is this concept useful? Consider the same question, when does $Ax = b$ have a solution? We already know the story for finite dimensions. There are finitely many relations for b in this case (b is in the column space of the matrix A). However, the same is true for the case that A is pseudo-invertible. We will show that if A is pseudo-invertible, then $Y = R_A \oplus Z$ where Z is finite dimensional. Then we only need to check that b has no component in Z.

Another Form: Given A, B linear maps from X to Y, we say that $A \sim B$ if $A - B$ is of finite rank. Then

- $\sim$ is an equivalence relation.
- If $A_1 \sim B_1, A_2 \sim B_2$, then $A_1 + A_2 \sim B_1 + B_2$
- If $A_1 \sim B_1, A_2 \sim B_2$, and $X \xrightarrow[B_1]{A_1} Y \xrightarrow[B_2]{A_2} Z$, then $A_2 A_1 \sim B_2 B_1$.

Fact: A: $X \rightarrow Y$ is pseudoinvertible if and only if there exists a linear map B from Y to X such that $AB \sim I$ and $BA \sim I$. Note that invertibility is the property above with $\sim$ replaced by $=$.

Fact: If $A: X \rightarrow Y$ linear such that $B_1 A \sim A B_2 \sim I$, then $B_1 \sim B_2$ and B_1 is a pseudoinverse of A.

Proof. $B_1 = B_1 I \sim B_1 A B_2 \sim I B_2 = B_2$, so $B_1 \sim B_2$. And since $A B_1 \sim A B_2 \sim I$, B_1 is a pseudoinverse of A. ☐

Proposition 131. *Let A: $X \rightarrow Y$ linear. Then A is pseudoinvertible if and only if* $\dim N_A < \infty$ *and* $\dim (Y / R_A) < \infty$.

Y / R_A is called the "corange" of A and its dimension is the "codimension" of A.

Before we prove the result, let's look at a few examples:

Example 132. Consider T: $l^2 \rightarrow l^2$ as before, the shift operator $T(a_1, a_2, ...) = (0, a_1, a_2, ...)$. Then T is pseudoinvertible, since if we define the reverse shift $T'(a_1, a_2, ...) = (a_2, ...)$, we see that $T' \circ T = I$, and

$$T(T'(a_1, a_2, ...)) = (0, a_2, ...)$$

so that $TT' = I - K$ where $K(a_1, a_2, ...) = (a_1, 0, 0, ...)$ has finite rank. Thus $TT' \sim T'T \sim I$.

Example 133. Let $X = C([0, 1])$ (continuous functions on $[0, 1]$). Then if we define a linear map T mapping $X \rightarrow X$ by

$$(Tf)(x) = f(x) + \int_0^1 e^{xy} f(y)\, dy$$

The equation $Tf = g$ comes up in practice, and we will discuss this later in the Fredholm theory.

We will show that T is pseudoinvertible later when we discuss compact operators. It turns out that the integral operator $f \mapsto \int_0^1 e^{xy} f(y)\,dy$ is a compact operator (Hilbert-Schmidt kernels, Example 149), and that an operator in the form $I + K$ with K compact is pseudoinvertible.

Before the proof of the Proposition 131, we introduce a general technique:

Block Decomposition: We use block matrices to organize computation. If we decompose $X = X_1 \oplus \ldots \oplus X_n$ and $Y = Y_1 \oplus \ldots \oplus Y_m$, then letting P_j be the projection of X to X_j and Π_i be the projection of Y to Y_i. We can define $A_{ij}: X_j \to Y_i$ by $A_{ij} = \Pi_i \circ A \circ P_j$ (project to X_j, then map A, then project to Y_i), and then $A = \sum_{ij} A_{ij}$. Furthermore, we can express this action in a matrix. Decomposing any x as $x = x_1 + \ldots + x_n$, we have

$$\begin{pmatrix} \Pi_1 A x \\ \vdots \\ \Pi_m A x \end{pmatrix} = \begin{pmatrix} A_{11} & \ldots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \ldots & A_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

This is a purely algebraic identity, and thus it works for infinite dimensional spaces as well (decomposing them into a direct sum of finitely many subspaces).

Observation: If $X_1 \subset X$ is a subspace, then there exists a subspace $X_2 \subset X$ such that $X = X_1 \oplus X_2$.

These two points will be used in the proof.

Proof. (of Proposition 131) First suppose that A is pseudoinvertible. This means there exists an operator B from $Y \to X$ such that $BA = I + K$ and $AB = I + L$ where K, L are finite rank. Note that

$$N_A \subset N_{BA} = N_{I+K}$$

Note that if $x \in N_{I+K}$ then $x + Kx = 0$, thus $x = K(-x) \in R_K$, and so $N_{I+K} \subset R_K$ and

$$\dim N_A \leq \dim N_{I+K} \leq \dim R_K < \infty$$

As for Y / R_A, we make a few observations:

- $R_A \supset R_{AB} = R_{I+L}$
- $Y / R_A \cong (Y / R_{AB}) / (R_A / R_{AB})$

  This is because we can write $R_A = R_{AB} \oplus W$, where $W \cong R_A / R_{AB}$, and so

  $$y + R_A \longleftrightarrow y + R_{AB} \oplus W \longleftrightarrow (y + R_{AB}) + R_A / R_{AB}$$

- $\dim Y / R_{I+L} < \infty$

  This is because if we consider the injection map Π from Y to Y / R_{I+L}, defined by

  $$\Pi(y) = y + R_{I+L}$$

  then we note that given $x + R_{I+L} \in Y / R_{I+L}$ we have that $-Lx \in R_L$ and

  $$\Pi(-Lx) = -Lx + R_{I+L} = -Lx + (I + L)x + R_{I+L} = x + R_{I+L}$$

  so that

  $$\Pi(R_L) = Y / R_{I+L}$$

and since $\dim R_L < \infty$, we have that $\dim(Y/R_{I+L}) < \infty$ as well.

Finally, we have that

$$\dim(Y/R_A) = \dim(Y/R_{AB}) - \dim(R_A/R_{AB}) \le \dim(Y/R_{AB}) = \dim(Y/R_{I+L}) < \infty$$

This shows that if A is pseudoinvertible, then $\dim(N_A)$ and $\dim(Y/R_A)$ are finite.

Conversely, suppose that $\dim(N_A)$ and $\dim(Y/R_A)$ are finite. Decompose $X = X_1 \oplus N_A$ and $Y = R_A \oplus Y_1$. Then we can write A in block matrix form

$$A = \begin{array}{c c} & \begin{array}{cc} X_1 & N_A \end{array} \\ \begin{array}{c} R_A \\ Y_1 \end{array} & \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} \end{array}$$

Then we claim that $A_1: X_1 \to R_A$ given by $A_1 = A|_{X_1}$ is invertible. Injectivity follows from the fact that for any $x_1 \in X_1$ such that $Ax_1 = 0$, we see that $x_1 \in X_1 \cap N_A = \{0\}$. Surjectivity follows from the fact that given $y \in R_A$ and x with $Ax = y$, we have that $x = x_1 + x_2$ with $x_1 \in X_1$ and $x_2 \in N_A$. In this case

$$y = Ax = Ax_1 = A_1 x_1$$

Then, we simply define

$$B = \begin{array}{c c} & \begin{array}{cc} R_A & Y_1 \end{array} \\ \begin{array}{c} X_1 \\ N_A \end{array} & \begin{pmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \end{array}$$

in which case

$$AB = \begin{array}{c c} & \begin{array}{cc} R_A & Y_1 \end{array} \\ \begin{array}{c} R_A \\ Y_1 \end{array} & \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \end{array} = I + \begin{pmatrix} 0 & 0 \\ 0 & -I|_{Y_1} \end{pmatrix}$$

where the last term above has finite rank, since $\dim(Y_1) = \dim Y - \dim R_A < \infty$. We can do the same for BA, and this shows that $AB \sim BA \sim I$, so A is pseudoinvertible. □

Definition 134. *If $A: X \to Y$ is pseudoinvertible, then we define its **index** to be*

$$\operatorname{ind} A = \dim N_A - \dim Y/R_A$$

Example 135. If $A: X \to Y$ where both spaces are finite dimensional, then

$$\operatorname{ind} A = \dim N_A - \dim Y + \dim R_A = \dim X - \dim Y$$

and in particular, if $A: X \to X$ with $\dim X < \infty$ then $\operatorname{ind} A = 0$.

Example 136. For the shift operator, $T: l^2 \to l^2$, $\dim N_T = 0$ and $\dim l^2/R_T = 1$, so the index is -1.

Proposition 137.

1. *If A, B are pseudoinvertible, then so is BA and* $\operatorname{ind} BA = \operatorname{ind} B + \operatorname{ind} A$.

2. *If A is pseudoinvertible and K is finite rank, then* $\operatorname{ind}(A + K) = \operatorname{ind} A$

Proof.

(1) exact sequence $0 \to N_A \to N_{BA} \to N_B \to Y/R_A \to Z/R_{BA} \to Z/R_B \to 0$

(2) first prove for $I + K$ using decomposition $X = X_1 \oplus X_2$, with $\dim X_1 < \infty$, $X_2 \subset N_K$ and $R_K \subset X_1$.

More precisely, $X = (R_K \oplus X_4) \oplus (N_K \cap X_3)$. □

Fredholm Operators

Let X, Y now be Banach spaces. If $A \in \mathcal{L}(X, Y)$ such that $\dim N_A < \infty$, R_A is closed and $\dim Y/R_A < \infty$, then we say that A is a **Fredholm operator**. We define the **index** of such an operator to be

$$\operatorname{ind}(A) = \dim N_A - \dim Y/R_A$$

as with pseudoinvertible operators. Note that we have Banach space structure here, and Fredholm operators have the additional condition that R_A is closed, though actually this is not an additional restriction. It turns out that

Proposition 138. *Let $A \in \mathcal{L}(X, Y)$. Then*

$$\begin{aligned} A \text{ is Fredholm} &\iff \dim N_A < \infty, \dim Y/R_A < \infty, R_A \text{ closed} \\ &\iff \dim N_A < \infty, \dim Y/R_A < \infty, \\ &\iff A \text{ is pseudoinvertible} \end{aligned}$$

We have already proved the last equivalence, and the first equivalence is a definition. The middle follows from open mapping theorem. We will show that if $\dim Y/R_A < \infty$, then R_A is closed. The reason is that first, without loss of generality $N_A = \{0\}$, otherwise we replace A with $\overline{A}$: $X/N_A \to Y$ with $\overline{A}(x + N_A) = A\,x$. Then we have $Y = R_A \oplus Y_1$ with $\dim Y_1 < \infty$. Then we can define a map φ: $X \oplus Y_1 \to Y$ by

$$\varphi(x, \xi) = A\,x + \xi$$

and since $Y = R_A \oplus Y_1$, this is a bijection. φ is one-to-one since A is one-to-one, and given any $y \in Y$, we can decompose $y = A\,x + \xi$ for some x, ξ. By open mapping theorem, φ is a homeomorphism, and since $R_A = \varphi(X)$, $\varphi(X)$ is closed and therefore R_A is closed.

In addition to the pseudoinvertible property that A, B Fredholm implies $B\,A$ is Fredholm with

$$\operatorname{ind}(B\,A) = \operatorname{ind} B + \operatorname{ind} A$$

we also have the following special property. Denote $\operatorname{Fred}(X, Y) \subset \mathcal{L}(X, Y)$ the subspace of operators that are Fredholm.

Proposition 139. $\operatorname{Fred}(X, Y) \subset \mathcal{L}(X, Y)$ *is an open subset, and specifically,*

$$\operatorname{ind}(A + E) = \operatorname{ind} A$$

for all $|E| < \varepsilon$ with ε sufficiently small.

Proof. Let $A \in \operatorname{Fred}(X, Y)$. We will use block decomposition again. Let $X = X_1 \oplus N_A$, with X_1 closed and N_A finite dimensional (recall finite dimensional subspaces have a closed complement). Also, $Y = R_A \oplus Y_1$, with Y_1 finite dimensional, R_A closed since A is Fredholm.

Now

$$A = \begin{array}{c} \begin{array}{cc} X_1 & N_A \end{array} \\ \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ R_A \\ Y_1 \end{array}$$

where $A_1 = A|_{X_1}$. Note that A_1: $X_1 \to R_A$ is a bijection and thus by open mapping A_1 is a homeomorphism. Now we write

$$A + E = \begin{array}{c} \begin{array}{cc} X_1 & N_A \end{array} \\ \begin{pmatrix} A_1 + E_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix} \end{array} \begin{array}{c} \\ R_A \\ Y_1 \end{array}$$

for $|E| < \varepsilon$, E_{ij} are the appropriate restrictions of E, i.e. $E_{11} = \pi_{R_A} E|_{X_1}$. For ε sufficiently small, we note that $A_1 + E_{11}$ is invertible, more specifically, when $|E_{11}| < \frac{1}{|A_1^{-1}|}$. This is because we can write

$$A_1 + E_{11} = A_1(I + A_1^{-1} E_{11})$$

and use the Neumann series $I - B = \sum_i B^i$ when $|B| < 1$. Thus, we now multiply $A + E$ by invertible operators and use the property that the composition of psueodinvertible operators is also pseudoinvertible. Computation gives

$$\begin{array}{c} \begin{array}{cc} R_A & Y_1 \end{array} \\ \begin{pmatrix} I & 0 \\ -E_{21}(A_1 + E_{11})^{-1} & I \end{pmatrix} \end{array} \begin{array}{c} \\ R_A \\ Y_1 \end{array} \begin{array}{c} \begin{array}{cc} X_1 & N_A \end{array} \\ \begin{pmatrix} A_1 + E_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix} \end{array} \begin{array}{c} \\ R_A \\ Y_1 \end{array} \begin{array}{c} \begin{array}{cc} X_1 & N_A \end{array} \\ \begin{pmatrix} I & -(A_1 + E_{11})^{-1} E_{12} \\ 0 & I \end{pmatrix} \end{array} \begin{array}{c} \\ X_1 \\ N_A \end{array}$$

which reduces to

$$C = \begin{array}{c} \begin{array}{cc} X_1 & N_A \end{array} \\ \begin{pmatrix} A_1 + E_{11} & 0 \\ 0 & B \end{pmatrix} \end{array} \begin{array}{c} \\ R_A \\ Y_1 \end{array}$$

where $B = E_{22} - E_{21}(A_1 + E_{11})^{-1} E_{12}$. What's important here is that $A_1 + E_{11}$ is invertible and B is finite dimensional, and thus this operator C is pseudoinvertible. This implies that $A + E$ is also pseudoinvertible after reversing the procedure. Furthermore, we can compute the index. Since the index of invertible operators is 0, we note that $\operatorname{ind}(A + E) = \operatorname{ind}(C) = \operatorname{ind}(A_1 + E_{11}) + \operatorname{ind}(B)$ (this equality follows by inspection; for instance, the null space is the null space of $A_1 + E_{11}$ plus the null space of B). Since B is a finite dimensional operator, by Example 135 $\operatorname{ind}(B) = \dim N_A - \dim Y_1$, and since $A_1 + E_{11}$ is invertible its index is 0. Thus $\operatorname{ind}(A + E) = \dim N_A - \dim Y_1 = \operatorname{ind}(A)$.

□

Example 140. If $t \in [0, 1]$, then if $A(t)$: $[0, 1] \to \operatorname{Fred}(X, Y)$ is continuous, then $\operatorname{ind} A(s) = \operatorname{ind} A(t)$ for all $s, t \in [0, 1]$. This is referred to as "homotopy invariance of index".

Compact Operators

Recall in finite dimensions, for every operator (matrix) A, the spectrum consists entirely of eigenvalues (recall that in finite dimensions injectivity implies surjectivity). Compact operators in infinite dimensions enjoy a similar property.

Let X, Y be Banach spaces. If $A \in \mathcal{L}(X, Y)$ (bounded linear operator) such that

$$A B_1 = \{A x \colon |x| < 1\}$$

is precompact in Y, then we say that A is a **compact operator**.

Here *precompact* means that the closure is compact in the context of a complete metric space. For instance, in $\mathbb{R}^n$ all bounded sets are precompact.

Proposition 141. *Let X be a complete metric space, and let $A \subset X$. Then*

$$\begin{aligned} A \text{ precompact} &\iff \text{every sequence } x_j \in A \text{ has a Cauchy subsequence} \\ &\iff \text{for all } \varepsilon > 0, A = \bigcup_{i=1}^{m} A_i,\ \mathrm{diam}(A_i) < \varepsilon \\ &\iff \text{for all } \varepsilon > 0,\ \text{exists a finite } \varepsilon\text{-net} \end{aligned}$$

We have the following connection to pseudoinvertible operators.

Theorem 142. *Let $X/\mathbb{C}$ and $K \in \mathcal{L}(X)$ be compact. Then*

1. $\dim(N_{I-K}) < \infty$.
2. *R_{I-K} is closed*
3. $\dim X/R_{I-K} = \dim N_{I-K} < \infty$ *i.e. if $N_{I-K} = \{0\}$ then $X = R_{I-K}$.*

For proving Theorem 142, we will be making use of the following facts:

Proposition 143.

1. *If K_1, K_2 compact, then $K_1 + K_2$ is also compact.*
2. *If K is compact and A is bounded, then $K A$ is compact (also $A K$ is compact)*
3. *If K_i is a sequence of compact operators, and $K_i \to K$ (in operator norm), then K is compact.*
4. *If K is compact, K' is also compact.*

Proof. (1) We can find A_i with $\mathrm{diam}(A_i) < \varepsilon$, $K_1 B_1 = \bigcup_{i=1}^m A_i$ and likewise $K_2 B_2 = \bigcup_{j=1}^n C_j$ with $\mathrm{diam}(C_j) < \varepsilon$. Then

$$(K_1 + K_2) B_1 \subset K_1 B_1 + K_2 B_2 \subset \bigcup_{i,j} A_i + B_j$$

where $\mathrm{diam}(A_i + B_j) < 2\varepsilon$ (i.e. a finite 2ε-net)

(2) The proof uses the same idea. Let $|A| = M$. We have $K B_1 = \bigcup_{i=1}^m C_i$ with $\mathrm{diam}(C_i) < \varepsilon$. Then note $K A\ B_1 \subset K B_M = \bigcup_{i=1}^m M C_i$ where $\mathrm{diam}(M\ C_i) \leq M\varepsilon$. We also have $A\ K\ B_1 \subset \bigcup_{i=1}^m A\ C_i$, and $\mathrm{diam}(A\, C_i) \leq M\,\varepsilon$ also ($|A\,x - A\,y| \leq M\,|x - y|$).

(3) The idea here is to take K_i with $|K_i - K| < \varepsilon$, and use $K_i B_1 = \bigcup_{j=1}^m A_j$ with $\mathrm{diam}(A_j) < \varepsilon$, with

$$K B_1 \subset K_i B_1 + (K - K_i) B_1 \subset \bigcup_j A_j + (K - K_i) B_1$$

where $\mathrm{diam}(A_j + (K - K_i) B_1) < 2\varepsilon$, so we have a finite 2ε net for $K B_1$.

(4) Recall that $K': Y' \to X'$ is defined by $K' l(x) = l(K x)$. Examine

$$\begin{aligned} K' B_1^{Y'} &= \{K' l, |l| < 1\} \\ &= \{l \circ K, |l| < 1\} \end{aligned}$$

Let $F = \overline{K B_1}$, which is a compact subset of Y. We will show the precompactness of $K' B_1^{Y'}$. Note given $l_1, l_2 \in Y'$, we have

$$\begin{aligned} |l_1 \circ K - l_2 \circ K|_{X'} &= \sup_{|x| \le 1} |l_1 K x - l_2 K x| \\ &= \sup_{y \in F} |l_1(y) - l_2(y)| \end{aligned}$$

The result will follow from Arzela Ascoli, using $C(F, \mathbb{C})$ and the compactness of F. Note that

$$\left\{ l|_F, |l| < 1 \right\} \subset C(F, \mathbb{C})$$

we have $l|_F$ uniformly Lipchitz (with constant 1) on a compact set F, and thus this family of functions is equicontinuous. Given a bounded sequence $l_i \circ K \in K' B_1^{Y'}$, note that $l_i|_F$ is an equicontinuous sequence, and by Arzela Ascoli there exists a subsequence $l_{i'}|_F$ which is Cauchy uniformly, i.e.

$$\sup_{y \in F} |l_{i'}(y) - l_{j'}(y)| \to 0$$

which by above implies that $|l_{i'} \circ K - l_{j'} \circ K|_{X'} \to 0$, and hence $K' B_1^{Y'}$ is precompact and K' is compact.

□

Proof. (of Theorem 142) Let $A = I - K$.

(1) Here we show that $\dim N_A < \infty$. Suppose that $\dim N_A = \infty$. Then there exists $x_j \in N_A$ such that $|x_j| = 1$ and $|x_j - x_k| \ge \frac{1}{2}$ for $1 \le k < j$. Since K is compact, there exists a subsequence $K x_{n'}$ which is Cauchy. Then on one hand, we have that $|K x_{n'} - K x_{m'}| \to 0$ but on the other, $|x_{n'} - x_{m'}| \ge \frac{1}{2}$, and this is a contradiction since K is bounded.

(2) Now we show that R_A is closed. Recall that this is implied by $\dim Y / R_A < \infty$, but we will show this directly (without directly appealing to open mapping). Write $X = N_A \oplus X_1$ where X_1 is closed. The claim is that $|A x| \ge c|x|$ for $x \in X_1$. Then this implies that $R_A = A X_1$ is closed, since under this condition if $A x_j$ is Cauchy, then x_j is Cauchy, and by continuity the limit of x_j maps to the limit of $A x_j$.

To prove the inequality, suppose it is false, then there exists a sequence $x_j \in X_1$ where $|A x_j| < \frac{1}{j}|x_j|$. By scaling, we can take $|x_j| = 1$, and then $A x_j \to 0$, which means that $x_j - K x_j \to 0$. Passing to a subsequence, by compactness of K we can find $K x_{j'} \to y$, and then $x_{j'} \to y$. This implies that $K y \to y$ and $y \in N_A$. But since X_1 is closed, $y \in X_1$ also. Thus $y = 0$. However, we have chosen $x_{j'}$ so that $|y| = 1$, and this is a contradiction.

(3) Next we show $\dim X / R_A < \infty$, and that $\dim X / R_A = \dim N_A$. Note that $A' = I - K'$ is compact by Proposition 143 (4). By above, $\dim N_{A'} < \infty$. We will use the isometries $(X / R_A)' \cong (R_A)^{\perp} = N_{A'},$. Then

$$\dim (X / R_A)' = \dim N_{A'} < \infty$$

and since $(X / R_A)'$ is a finite dimensional Banach space, $\dim X / R_A = \dim (X / R_A)' < \infty$.

The identification is, given $l: (R_A)^{\perp}$, define $l'(x+R_A) = l(x)$ (well defined since l annihilates R_A), and we can go backwards from $l' \in (X/R_A)'$ as well. This is a norm preserving bijection also, since

$$\begin{aligned} |l|_{X'} &= \sup_{|x| \leq 1} |l(x)| \\ &= \sup_{\substack{|x| \leq 1 \\ y \in R_A}} |l(x+y)| \\ &= \sup_{|x+R_A| \leq 1} l'(x+R_A) \\ &= |l'|_{(X/R_A)'} \end{aligned}$$

recalling that $|x+R_A| = \inf_{y \in R_A} |x+y|$.

Now consider the map $A(t)$: $[0,1] \to \mathrm{Fred}(X,Y)$ given by $A(t) = I - t\,K$. Note that $A(0) = I$ and $A(1) = I - K$. By homotopy invariance, we have that $\mathrm{ind}(I-K) = \mathrm{ind}(I) = 0$, and this implies that

$$\dim N_A = \dim X/R_A$$

and as a consequence, $I-K$ is injective if and only if $I-K$ is surjective. $\square$

Proposition 144. *For K compact, we also have*

$$N_{I-K} \subset \cdots \subset N_{(I-K)^m} = N_{(I-K)^{m+1}} = \cdots$$

i.e. $N_{(I-K)^m} = N_{(I-K)^{m+1}}$ for some m.

Remark 145. This property is easy to see in finite dimensions, for instance the matrix

$$I-K = \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

where $N_{I-K} \subset N_{(I-K)^2} = N_{(I-K)^3} = \cdots$

Proof. Suppose that this property does not hold, and let $A = I-K$. This means that there exists x_n with $|x_n| = 1$, $(I-K)^n x_n = 0$ and $d(x_n, N_{(I-K)^{n-1}}) \geq \frac{1}{2}$. (each $x_n \in N_{(I-K)^n} \backslash N_{(I-K)^{n-1}}$).

Note that $Kx_n = x_n + y_n$ where $y_n \in N_{(I-K)^{n-1}}$. Then, by compactness of K, we can find a subsequence $Kx_{n'}$ such that $|Kx_{n'} - Kx_{m'}| \to 0$. On the other hand, we have that

$$\begin{aligned} |Kx_{n'} - Kx_{m'}| &= |x_{n'} + y_{n'} - x_{m'} - y_{m'}| \\ &= |x_{n'} - (y_{n'} + x_{m'} + y_{m'})| \\ &\geq \frac{1}{2} \end{aligned}$$

since $y_{n'} + x_{m'} + y_{m'} \in N_{(I-K)^{n'-1}}$ (if $n' > m'$), and this is a contradiction. $\square$

Alternate Characterization of Fredholm Operators.

Proposition 146. *Let $A: \mathcal{L}(X,Y)$. If there exist $B, C \in \mathcal{L}(Y,X)$ such that $A\,B = I + K$ and $C\,A = I + L$ with K, L compact, then A is a Fredholm operator.*

Remark 147. This is weaker than the condition for pseudoinvertibility. Recall that for pseudoinvertibility we need a *single* operator $B \in \mathcal{L}(Y, X)$ so that $A\,B = I + K$, $B\,A = I + L$ with K, L *finite rank*. Thus, note that if A is a Fredholm operator, then this condition certainly holds (since A Fredholm $\Longleftrightarrow A$ pseudoinvertible).

Proof. Note that $N_A \subset N_{CA} = N_{I+L}$ which is finite dimensional. Thus $\dim N_A < \infty$. Also we have that $R_A \supset R_{AB} = R_{I+K}$, and this implies that $\dim Y/R_A \leq \dim Y/R_{I+K} < \infty$. Thus by the equivalences in Proposition 138, A is Fredholm. $\square$

Corollary 148. *If $A: X \to Y$ is Fredholm and $K: X \to Y$ is compact, then $A + K$ is Fredholm with*

$$\mathrm{ind}(A+K) = \mathrm{ind}(A)$$

Proof. Since A is Fredholm and thus pseudoinvertible, we can find $B: Y \to X$ such that $A\,B = I + K_1$ and $B\,A = I + K_2$ where K_1, K_2 are finite rank. Then we have that

$$\begin{aligned} (A+K)B &= I + K_1 + K\,B \\ B(A+K) &= I + K_2 + BK \end{aligned}$$

and since $K_1 + K\,B, K_2 + B\,K$ are compact (Proposition 143), $A + K$ is Fredholm. Furthermore, by homotopy invariance, we can take $A(t) = A + t\,K$ which is continuous operator from $[0, 1] \to \mathrm{Fred}(X, Y)$, and thus

$$\mathrm{ind}(A) = \mathrm{ind}\,A(0) = \mathrm{ind}\,A(1) = \mathrm{ind}(A+K)$$

$\square$

Example 149. *Hilbert-Schmidt Kernels.* Let (X, μ) and (Y, ν) be measure spaces, let $K(x, y) \in L^2(X, Y)$ in the product space $L^2(X, Y) = L^2((X \times Y), \mu \times \nu)$, and define $T_K: L^2(X) \to L^2(Y)$ by

$$T_K f(y) = \int_X K(x, y)\, f(x)\, dx$$

Then T_K is a compact operator.

Proof. First we show that T_K is bounded. This is simply by Hölder:

$$\begin{aligned} |T_K f(y)|^2 &\leq \left(\int_X |K(x, y)|^2\, dx\right) |f|^2_{L^2(X)} \\ |T_K f(y)|^2_{L^2(Y)} = \int_Y |T_K f(y)|^2 dy &\leq |K(x, y)|^2_{L^2(X \times Y)} |f|^2_{L^2(X)} \end{aligned}$$

and thus the operator norm is $|T_K|_{\mathcal{L}(L^2(X), L^2(Y))} \leq |K|_{L^2(X,Y)}$.

To show that T_K is compact, we will find a sequence of finite rank operators $T_{K_n} \to T_K$ convergent in operator norm, and since finite rank operators are compact, by Proposition 143 T_K will be compact.

Using an orthonormal basis for $L^2(X \times Y)$, we can write $K(x, y) = \sum_{i=1}^{\infty} p_i(x) \psi_i(y)$ where $p_i \in L^2(X)$ and $\psi_i(y) \in L^2(Y)$ (this is accomplished by fixing x and expanding K in terms of a basis for $L^2(Y)$, and then expanding each coefficient in terms of a basis for $L^2(X)$). Now if we truncate the series

$$K_n(x, y) = \sum_{i=1}^{n} p_i(x) \psi_i(y)$$

and define the operator T_{K_n} by

$$T_{K_n} f(y) = \int_X K_n(x, y)\, f(x)\, dx = \sum_{i=1}^{n} \left(\int p_i(x)\, f(x)\, dx \right) \psi_i(y)$$

which is a finite linear combination and thus T_{K_n} is finite rank (and thus compact).

Note that $K_n \to K$ in $L^2(X \times Y)$, and by the same inequality used to show boundedness of T_K, this implies that $T_{K_n} \to T_K$ in operator norm, and therefore T_K is also compact. □

Spectrum of Compact Operators

Theorem 150. *Let $X/\mathbb{C}$ be a Banach space, $K \in \mathcal{L}(X)$ compact. Then*

1. *$\sigma(K)$ is countable.*

2. *$\lambda \in \sigma(K) \backslash \{0\} \Longrightarrow \lambda \in \sigma_p(K)$ (i.e. an eigenvalue), $\dim N_{\lambda I - K} < \infty$ and*

$$N_{\lambda I - K} \subset N_{(\lambda I - K)^2} \subset \cdots \subset N_{(\lambda I - K)^m} = N_{(\lambda I - K)^{m+1}} = \cdots$$

3. *$\sigma(K) \backslash \{0\}$ has no nonzero limit point (the spectrum is isolated)*

Proof. First note that (3) implies (1). First we prove (2). Let $\lambda \in \sigma(K) \backslash \{0\}$ This means that $\lambda I - K$ is not invertible. Note that for Fredholm operators, injectivity is equivalent to surjectivity, so therefore $N_{\lambda I - K} \neq \{0\}$ and $\lambda \in \sigma_p(K)$. Note that $\dim N_{\lambda I - K} < \infty$ since $\lambda I - K$ is Fredholm, and we proved that the generalized eigenspace $\bigcup_{m=1}^{\infty} N_{(\lambda I - K)^m}$ is finite dimensional in Proposition 144.

Next we prove (3). Suppose there exists a nonzero limit point in $\sigma(K) \backslash \{0\}$. Then we have a sequence of eigenvalues $\lambda_i \in \sigma_p(K) \backslash \{0\}$, $\lambda_i \neq \lambda_j$ for $i \neq j$ with $\lambda_i \to \lambda \neq 0$. Let $X_i = N_{\lambda_i I - K}$. Note that X_i is invariant under K, i.e. $K X_i \subset X_i$,, since $\lambda_i x = K x \Longrightarrow \lambda_i K x = K^2 x$.

Also, $\operatorname{span}\left\{ \bigcup_i X_i \right\} = \bigoplus_i X_i$ since eigenvectors with different eigenvalues are linearly independent. Now let $Y_n = \bigcup_{i=1}^{n} X_i$ so that $Y_1 \subsetneq Y_2 \subsetneq Y_2 \subsetneq \cdots$, and choose y_k such that $|y_k| = 1$ and $d(y_k, Y_{k-2}) \geq \frac{1}{2}$. Since K is compact, there is a Cauchy subsequence $K y_{n'}$. Note that $y_{n'} = x_{n'} + z_{n'}$ with $x_{n'} \in X_{n'}$ and $z_{n'} \in Y_{n'-1}$, and that

$$\begin{aligned} K y_{n'} &= K x_{n'} + K z_{n'} \\ &= \lambda_{n'} x_{n'} + K z_{n'} \end{aligned}$$

where $K z_{n'} \in Y_{n'-1}$. Thus, on one hand by Cauchy property $|K y_{n'} - K y_{m'}| \to 0$, but on the other hand,

$$|K y_{n'} - K y_{m'}| = |\lambda_{n'}| \left| x_{n'} + \frac{1}{\lambda_{n'}} K z_{n'} \right| \geq \frac{1}{2} |\lambda_{n'}| \longrightarrow \frac{1}{2} |\lambda| \neq 0$$

which is a contradiction.

□

Spectrum of Compact Operators in Hilbert Spaces

Now we let H be a complex Hilbert space. Recall that the adjoint A^* of $A \in \mathcal{L}(H)$ is defined by

$$\langle A^* x, y \rangle = \langle x, A y \rangle$$

for all $x, y \in H$.

Recall in linear algebra:

1. If $A^* = A$ (i.e. self-adjoint), then $A = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}U^*$ with U unitary and λ_i real.

2. If $A^*A = AA^*$ (i.e. normal), then $A = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}U^*$ with U unitary and λ_i complex.

We will be proving the analogous result for compact self-adjoint and normal operators:

Theorem 151. (Spectral Theorem for Compact Self-Adjoint Operators) *If H is a separable Hilbert space, $K \in \mathcal{L}(H)$ compact and self-adjoint, then there exists an orthonormal basis $(u_j)_{j=1}^{\infty}$ such that $Ku_j = \lambda_j u_j$, $\lambda_j \in \mathbb{R}$ and $\lambda_j \to 0$.*

Theorem 152. (Spectral Theorem for Compact Normal Operators) *Let $K \in \mathcal{L}(H)$ be a normal compact operator, then there exists an orthonormal basis $(u_j)_{j=1}^{\infty}$ such that $Ku_j = \lambda_j\, u_j$, $\lambda_j \in \mathbb{C}$ and $\lambda_j \to 0$.*

Note the following characterization of self-adjoint operators on *complex* Hilbert spaces:

Proposition 153. *If $A \in \mathcal{L}(H)$, then*

$$\begin{aligned} A^* = A &\iff \langle Ax, y\rangle = \langle x, Ay\rangle \text{ for all } x, y \in H \\ &\iff \langle Ax, x\rangle \in \mathbb{R} \text{ for all } x \in H \end{aligned}$$

Proof. The first equivalence is by definition. Given that $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in H$, then we have that $\langle Ax, x\rangle = \langle x, Ax\rangle$ and thus $\langle Ax, x\rangle \in \mathbb{R}$. Given that $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in H$, note

$$\langle A(x - iy), x - iy\rangle = \langle Ax, x\rangle + i\langle Ax, y\rangle - i\langle Ay, x\rangle + \langle Ay, y\rangle \in \mathbb{R}$$

This implies that $\langle Ax, y\rangle - \langle Ay, x\rangle = 0$. $\square$

Definition: If $A \in \mathcal{L}(H)$, $A^* = A$ and $\langle Ax, x\rangle \geq 0$ for all $x \in H$, then we use the notation $A \geq 0$ (A is non-negative definite).

Proposition 154. *If $A \geq 0$, then $|A| = \sup_{|x|=1} \langle Ax, x\rangle$.*

Proof. Cauchy Schwarz shows $\sup_{|x|=1} \langle Ax, x\rangle \leq |A|$. For the other direction, let $M = \sup_{|x|=1} \langle Ax, x\rangle$. Note that $\langle Ax, y\rangle + \varepsilon\langle x, y\rangle$ is an inner product on H (the ε is necessarily to make a strictly positive quadratic form $\langle Ax, x, \rangle + \varepsilon\langle x, x\rangle > 0$. Then Cauchy Schwarz gives

$$|\langle Ax, y\rangle + \varepsilon\langle x, y\rangle| \leq \sqrt{\langle Ax, x\rangle + \varepsilon|x|^2}\,\sqrt{\langle Ay, y\rangle + \varepsilon|y|^2}$$

Letting $\varepsilon \to 0$, we have

$$\begin{aligned} |\langle Ax, y\rangle| &\leq \sqrt{\langle Ax, x\rangle} \cdot \sqrt{\langle Ay, y\rangle} \\ &\leq M\,|x||y| \end{aligned}$$

for all x, y. Taking supremum over x, y shows that $|A| \leq M$ (recall that $|x| = \sup_y |\langle x, y\rangle|$). $\square$

Corollary 155. *For any $A \in \mathcal{L}(H)$, $|A^*A| = |A|^2$.*

Proof.

$$\begin{aligned}|A^*A| &= \sup_{|x|=1} \langle A^*Ax, x\rangle\\ &= \sup_{|x|=1} \langle Ax, Ax\rangle\\ &= \sup_{|x|=1} |A^2x|\\ &= |A^2|\end{aligned}$$

□

As a direct consequence, we have the following:

Corollary 156. *If A is normal, then $r_A = |A|$*

Proof. We have that

$$\begin{aligned}|A^2|^2 &= |(A^2)^*A^2| = |(A^*A)^2|\\ &= |A^*A|^2 = |A|^4\end{aligned}$$

where we have applied the previous two results. Since $A^*A \geq 0$, by Proposition 154,

$$|(A^*A)^2| = \sup_{|x|=1} \left\langle (A^*A)^2 x, x\right\rangle = \sup_{|x|=1} |A^*Ax|^2 = |A^*A|^2$$

Thus we have that $|A^2| = |A|^2$, and repeating the argument, we have that $|A^{2^n}| = |A|^{2^n}$. Thus, using the formula for spectral radius (Theorem 125) we have that

$$r_A = \lim_{n \to \infty} |A^n|^{1/n} = \lim_{n \to \infty} |A^{2^n}|^{1/2^n} = |A|$$

□

Remark 157. Combining the results above, if K is a compact, self-adjoint, non-negative definite operator $A \geq 0$, then

$$\lambda_1 = r_K = |K| = \sup_{|x| \leq 1} \langle Kx, x\rangle$$

(just like in finite dimensions). This also gives a hint for how to find eigenvalues for K that is not necessarily non-negative.

The key idea is that this quantity $\sup_{|x|=1} \langle Kx, x\rangle$ (no loss in changing $|x| \leq 1$ with $|x| = 1$ by linearity) yields both eigenvalues and eigenvectors. It is directly related to the Rayleigh quotient $R_K(x) = \frac{\langle Kx, x\rangle}{\langle x, x\rangle}$.

Proposition 158. *Let K be a compact, self-adjoint operator. Then*

$$\lambda = \sup_{|x|=1} \langle Kx, x\rangle$$

is an eigenvalue, the supremum is achieved by some vector z, and $Kz = \lambda z$, i.e. z is an eigenvector of K with eigenvalue λ.

Proof. (Following Lax, Ch 28) The idea is that by definition of sup, there is a sequence x_k with $|x_k| = 1$ and $\langle Ax_k, x_k\rangle \to \lambda$. Since x_k is a bounded sequence in a Hilbert space, we can replace x_k with a convergent subsequence so that $x_k \to z$ weakly.

Now we claim that the compactness of A implies that $Ax_k \to Az$ in norm. Note that $Ax_k \to Az$ weakly since $l \circ A$ is a bounded linear functional so that $x_k \to z$ weakly implies that $l(Ax_k) \to l(Az)$ for all $l \in X'$. We already know that by compactness we can replace x_k with a further subsequence so that Ax_k converges in norm to some y. But this implies $Ax_k \to y$ weakly as well, and by uniqueness of limits of weak convergence, $y = Az$ and $Ax_k \to Az$ in norm.

Now we have that $\langle Ax_k, x_k\rangle \to \langle Az, z\rangle$ since

$$\begin{aligned}
|\langle Ax_k, x_k\rangle - \langle Az, z\rangle| &\leq |\langle Ax_k - Az, x_k\rangle| + |\langle Az, x_k\rangle - \langle Az, z\rangle| \\
&\leq |Ax_k - Az||x_k| + |\langle Az, x_k\rangle - \langle Az, z\rangle| \\
&\to 0
\end{aligned}$$

noting that $|Ax_k - Az| \to 0$ by norm convergence and $y \mapsto \langle Az, y\rangle$ is a linear functional so the second term vanishes by weak convergence of $x_k \to z$.

Thus, $\langle Az, z\rangle = \lambda_1$. Note that by lower semicontinuity of the norm, $|z| \leq 1$, but in fact z is also a unit vector, since otherwise $\left\langle A\frac{z}{|z|}, \frac{z}{|z|}\right\rangle = \frac{\lambda_1}{|z|^2} > \lambda_1$ which contradicts λ_1 being the supremum of $\langle Ax, x\rangle$.

The rest is just like in linear algebra: to show that $Az = \lambda_1 z$, note that z maximizes the Rayleigh quotient $R_A(x) = \frac{\langle Ax, x\rangle}{|x|^2}$ (over all $x \neq 0$). Then taking an arbitrary vector w and $f(t) = R_A(z + tw)$ maps $\mathbb{R} \to H$ and achieves its maximum at $t = 0$, thus its derivative vanishes at $t = 0$. The derivative at $t = 0$ is

$$\begin{aligned}
\lim_{h\to 0} \frac{f(h) - f(0)}{h} &= \lim_{h\to 0} \frac{1}{h}\left[\frac{\langle A(z + hw), z + hw\rangle}{|z + hw|^2} - \frac{\langle Az, z\rangle}{|z|^2}\right] \\
&= \lim_{h\to 0} \frac{1}{h}\left[\frac{h\langle Az, w\rangle + h\langle Aw, z\rangle + h^2\langle Aw, w\rangle}{|z + hw|^2} + \langle Az, z\rangle\left(\frac{1}{|z + hw|^2} - \frac{1}{|z|^2}\right)\right] \\
&= \frac{\langle Az, w\rangle + \langle Aw, z\rangle}{|z|^2} + \lambda_1 \lim_{h\to 0} \frac{\langle z, z\rangle - \langle z + hw, z + hw\rangle}{|z|^2|z + hw|^2} \\
&= \frac{\langle Az, w\rangle + \langle Aw, z\rangle}{|z|^2} - \lambda_1 \frac{\langle w, z\rangle + \langle z, w\rangle}{|z|^4}
\end{aligned}$$

Setting this to zero gives

$$\operatorname{Re}\langle Az - \lambda_1 z, w\rangle = 0$$

for all w. This implies $Az = \lambda_1 z$ (set $w = Az - \lambda_1 z$), and thus z is an eigenvector of A with eigenvalue λ_1. ☐

Remark 159. The procedure above works for $\inf_{|x|=1}\langle Kx, x\rangle$ as well, which will give the negative eigenvalues and eigenvectors.

Having shown how to get an eigenvector z, we can prove the main theorem by iterating the procedure on K restricted to the complement of span$\{z\}$. A technical point is that for an operator K not necessarily nonnegative, we need to consider both $\sup_{|x|\leq 1}\langle Kx, x\rangle$ and $\inf_{|x|\leq 1}\langle Kx, x\rangle$

Proof. (of Theorem 151) Essentially, the previous proposition does all the work. Before we do anything, first take an orthonormal basis of the null space $\ker K$ (recall this is finite dimensional), and set it aside. If $K=0$ then we are already done. For the rest of the proof we will then gather the other eigenvectors corresponding to nonzero eigenvalues, and do this in such a way that $|\lambda_n|$ decreases to 0.

If $K \neq 0$, then consider $\sup_{|x|=1} \langle Kx, x\rangle$ or $\inf_{|x|=1} \langle Kx, x\rangle$, whichever has larger magnitude, and denote this quantity by λ_1. As described above, we can find a unit eigenvector u_1 with $Ku_1 = \lambda_1 u_1$. Note that since K is self-adjoint, the orthogonal complement of the eigenvector u_1 is invariant under K, since given $v \in u_1^{\perp}$,

$$\langle Kv, u_1\rangle = \langle v, Ku_1\rangle = \lambda_1 \langle v, u_1\rangle = 0$$

so that $Kv \in u_1^{\perp}$ as well. We can then repeat the procedure with $K|_{u_1^{\perp}}$ (the operator restricted to the orthogonal complement of u_1). We will then find a second eigenvalue λ_2 with $|\lambda_2| \leq |\lambda_1|$, and a corresponding eigenvector u_2. Continuing this procedure, we generate a sequence of eigenvectors u_n corresponding to eigenvalue λ_n and $\lambda_n \to 0$ (noting there is no cluster point other than 0). This grabs all nonzero eigenvectors since at each stage we always grab an eigenvector with the eigenvalue with largest magnitude.

Note that our collection of eigenvectors $(u_k)_{k=1}^{\infty}$ (now including the vectors in $\ker K$) is complete, since otherwise the orthogonal complement contains another eigenvector with nonzero eigenvalue, which is a contradiction since such an eigenvector would have been found in our procedure. $\square$

Remark 160. (Min-Max Description of Eigenvalues) Let A be compact and self-adjoint, and list the positive eigenvalues in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots$. Then as in finite dimensional linear algebra, there is a min-max principle for the n-th largest eigenvalue:

$$\begin{aligned}\lambda_k &= \inf_{\substack{V \subset H \\ \dim V = k-1}} \sup_{\substack{x \in V^{\perp} \\ |x|=1}} \langle Ax, x\rangle \\ &= \sup_{\substack{V \subset H \\ \dim V = k}} \inf_{\substack{x \in V \\ |x|=1}} \langle Ax, x\rangle\end{aligned}$$

The proof is similar to the same result in finite dimensions.

We sketch the method of proof for compact normal operators:

Proof. (Sketch of Theorem 152) The proof is similar to the finite dimensional spectral theory. We prove a result about commuting compact, self-adjoint, operators being simultaneously diagonalizable, or having the same eigenvectors. Then decompose a normal operator K into a self-adjoint operator A and an anti-self-adjoint operator B:

$$K = \underbrace{\left(\frac{K+K^*}{2}\right)}_{A} + \underbrace{\left(\frac{K-K^*}{2}\right)}_{B}$$

Since K is compact, K^* is also compact, and therefore A, B are compact. Since K is normal, A, B commute, and we can apply the result for commuting compact self-adjoint operators to A and iB, and this will give the eigenvectors for K, since if $Av = \lambda v$ and $iBv = \mu v$, then

$$Kv = (\lambda - i\mu)v$$

$\square$

Remark 161. Note that with the spectral theorem for a compact normal operator K, we can identify H with l^2 on which K becomes a multiplication operator:

$$\begin{array}{ccc} H & \xrightarrow{x\mapsto Kx} & H \\ \varphi\updownarrow & & \updownarrow\varphi \\ l^2 & \xrightarrow[x_k\mapsto\lambda_k x_k]{} & l^2 \end{array}$$

and the identification is $\varphi(x) = (\langle x, u_k\rangle)_{k=1}^{\infty}$.

Also, we can define for a bounded function f the operator

$$f(A)x = \sum_{k=1}^{\infty} f(\lambda_k)\,\langle x, u_k\rangle u_k$$

which is also bounded:

$$\|f(A)x\|_H = \|f\|_{\infty}\|x\|_H$$

What's missing: Spectral theorem for bounded operators in Hilbert spaces.