Functional Analysis: Lecture notes based on Folland

Johan Jonasson [*†‡]

October 2011

1 Normed vector spaces

A *vector space* (VS) consists of objects (such as vectors or functions) that can be added and multiplied by scalars in such a way that the commutative and distributive laws hold. The set of scalars can be any field K, but here we will always have $K = \mathbb{R}$ or $K = \mathbb{C}$.

Let $\mathcal{X}$ be a VS. A *norm* on $\mathcal{X}$ is a function $\|\cdot\| : \mathcal{X} \to [0, \infty)$ for which

- $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in K$ and $x \in \mathcal{X}$,
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathcal{X}$ (the triangle inequality),
- $\|x\| = 0$ iff $x = 0$.

A norm gives rise to a metric, $\rho(x, y) = \|x - y\|$, and hence to a topology, the so called *norm topology*. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are said to be *equivalent* if there exist $0 < C_1 < C_2 < \infty$ such that

$$\forall x \in \mathcal{X} : C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1.$$

Equivalent norms generate the same topology and the same Cauchy sequences.

A VS equipped with a norm is called a *normed vector space* (NVS).

[*] Chalmers University of Technology
[†] Göteborg University
[‡] jonasson@chalmers.se

A NVS which is also complete is called a *Banach space*.

The following theorem will provide an important tool for proving completeness of a NVS. Recall that for $\{x_n\} \subseteq \mathcal{X}$, we say that the series $\sum_n x_n$ is absolutely convergent if $\sum_n \|x_n\| < \infty$.

Theorem 1.1 *A NVS $\mathcal{X}$ is complete iff every absolutely convergent series converges (in norm).*

Proof. Suppose on one hand that $\mathcal{X}$ is complete and $\sum_n \|x_n\| < \infty$. Writing $S_N = \sum_1^N x_n$, we have

$$\|S_N - S_M\| = \|\sum_{N+1}^{M} x_n\| \leq \sum_{N+1}^{M} \|x_n\| \leq \sum_{N+1}^{\infty} \|x_n\| \to \infty$$

as $M, N \to \infty$. Thus $\{S_N\}$ is Cauchy and hence convergent.

On the other hand suppose that every absolutely convergent series converges and that $\|x_n - x_m\| \to 0$ as $m, n \to \infty$. Pick for each j the index n_j so that $m, n \geq n_j \Rightarrow \|x_m - x_n\| < 2^{-j}$. Setting $x_0 = 0$, we have

$$x_{n_k} = \sum_{1}^{k} (x_{n_j} - x_{n_{j-1}})$$

which has norm less than 1 by the triangle inequality. Hence x_{n_k} converges to some limit y. However for $n \geq n_k$ and k large enough,

$$\|x_n - y\| \leq \underbrace{\|x_n - x_{n_k}\|}_{<2^{-k}} + \underbrace{\|x_{n_k} - y\|}_{<2^{-k}} < 2^{-(k-1)}.$$

I.e. $x_n \to y$. $\square$

Example. Let X be a topological space and let $B(X)$ be the space of bounded continuous (complex-valued) functions on X. Define the norm $\|\cdot\|_u$ on $B(X)$ by

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

Suppose that $\sum_n \|f_n\|_u < \infty$. Then clearly

$$\|\sum_n f_n\|_u \leq \sum_n \|f_n\|_u,$$

so $f := \sum_n f_n \in B(X)$. Does $\sum_1^N f_n$ converge to f? Yes:

$$\|\sum_1^N f_n - f\|_u = \|\sum_{N+1}^{\infty} f_n\|_u \leq \sum_{N+1}^{\infty} \|f_n\|_u \to 0$$

as $N \to \infty$. By Theorem 1.1 we conclude that $B(X)$ is a Banach space. $\square$

Example. Let $(X, \mathcal{M}, \mu)$ be a measure space. We claim that $L^1(X, \mathcal{M}, \mu)$ is a Banach space. (Here we have to identify functions that are equal a.e.)

Recall that the norm is given by $\|f\|_1 = \int |f|$. Suppose that $\sum_n \|f_n\|_1 < \infty$. Then by the MCT,

$$\int \sum_n |f_n| = \sum_n \int |f_n| = \sum_n \|f_n\|_1 < \infty$$

so that $f := \sum_n f_n$ exists (a.e.) and, using the MCT again,

$$\|\sum_1^N f_n - f\|_1 = \int |\sum_{N+1}^{\infty} f_n| \leq \sum_{N+1}^{\infty} \int |f_n| = \sum_{N+1}^{\infty} \|f_n\|_1 \to 0$$

as $N \to \infty$. Now Theorem 1.1 proves our claim. $\square$

If $\mathcal{X}$ and $\mathcal{Y}$ are NVS's, then the *product norm* on $\mathcal{X} \times \mathcal{Y}$ is given by $\|(x, y)\| = \max(\|x\|, \|y\|)$. This norm is equivalent to $\|x\| + \|y\|$, $(\|x\|^2 + \|y\|^2)^{1/2}$, etc.

Next we consider *linear maps* from $\mathcal{X}$ to $\mathcal{Y}$. A linear map $T : \mathcal{X} \to \mathcal{Y}$ is said to be *bounded* if there exists $C < \infty$ such that $\|Tx\| \leq C\|x\|$. Here, of course, the norm notation refers to the norm of X for objects in $\mathcal{X}$ and to the norm of $\mathcal{Y}$ for objects in $\mathcal{Y}$. Note that this notion of boundedness is not the same as when we speak of a bounded function.

Proposition 1.2 *The following statements are equivalent*

(a) T *is bounded.*

(b) T *is continuous.*

(c) T *is continuous at 0.*

Proof. That (b) implies (c) is trivial. If (c) holds, then there exists a $\delta > 0$ such that $\|x\| < 2\delta \Rightarrow \|Tx\| < 1$. Thus, for any x,

$$\|Tx\| = \|\frac{\|x\|}{\delta}T(\frac{\delta x}{\|x\|})\| < \frac{1}{\delta}\|x\|,$$

i.e. T is bounded. Finally assume that (a) holds. Fix x and $\epsilon > 0$. Then if $\|x' - x\| < \epsilon/C$, we have

$$\|Tx' - Tx\| = \|T(x' - x)\| \leq C\|x' - x\| < \epsilon.$$

□

Let $L(\mathcal{X}, \mathcal{Y})$ be the space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$. In this notation it is also assumed that $L(\mathcal{X}, \mathcal{Y})$ is equipped with the *operator norm* given by

$$\|T\| = \sup\{\|Tx\| : \|x\| = 1\} = \sup\{\frac{\|Tx\|}{\|x\|} : x \in \mathcal{X}\}.$$

The space $L(\mathcal{X}, \mathcal{Y})$ is not always complete, but the following holds

Proposition 1.3 *If $\mathcal{Y}$ is complete, then so is $L(\mathcal{X}, \mathcal{Y})$.*

Proof. Assume that $\{T_n\} \subseteq L(\mathcal{X}, \mathcal{Y})$ is Cauchy. Then for arbitrary $x \in \mathcal{X}$,

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\|.$$

Hence $\{T_n x\}$ is Cauchy and hence convergent, since $\mathcal{Y}$ is Cauchy. We can thus define

$$Tx = \lim_n T_n x, x \in \mathcal{X}.$$

Then the operator T is linear and $\|T_n x\| \to \|Tx\|$ (see exercise 2), so

$$\|T\| = \sup\{\|Tx\| : \|x\| = 1\} = \sup\{\lim_n \|T_n x\| : \|x\| = 1\} \leq \limsup_n \|T_n\| < \infty$$

since $\{T_n\}$ is Cauchy. Hence $T \in L(\mathcal{X}, \mathcal{Y})$. We need to show that $\|T_n - T\| \to 0$. Pick $\epsilon > 0$ and N so that $m, n \geq N \Rightarrow \|T_n - T_m\| < \epsilon$. Then for any x with $\|x\| = 1$,

$$\|T_n x - Tx\| = \|T_n x - \lim_m T_m x\| = \lim_m \|T_n x - T_m x\| \leq \limsup_m \|T_n - T_m\| < \epsilon.$$

Hence $\|T_n - T\| < \epsilon$. □

Suppose that $T \in L(\mathcal{X}, \mathcal{Y})$ and $S \in L(\mathcal{Y}, \mathcal{Z})$. Then

$$\|S(T(x))\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|$$

so that $\|S \circ T\| \leq \|S\|\|T\|$ and in particular $S \circ T \in L(\mathcal{X}, \mathcal{Z})$.

If $T \in L(\mathcal{X}, \mathcal{Y})$ is bijective and T^{-1} is bounded (i.e. $T^{-1} \in L(\mathcal{Y}, \mathcal{X})$), then we say that T is an *isomorphism*. If T also has the property that $\|Tx\| = \|x\|$ for all x, then t is said to be an *isometry*. Whenever two NVS are isomorphic, the have corresponding notion of convergence. If they are also isometrically isomorphic, we may think of them as two interpretations of the same space. When $\mathcal{X}$ and $\mathcal{Y}$ are isometrically isomorphic, one writes for short $\mathcal{X} \cong \mathcal{Y}$.

2 L^p-spaces

Fix a measure space $(X, \mathcal{M}, \mu)$ and $1 \leq p < \infty$. For $f : X \to \mathbb{C}$ measurable, let

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

Then $L^p(X, \mathcal{M}, \mu)$ (or $L^p(X)$ or $L^p(\mu)$ or simply L^p when there is no risk for confusion) is the space of measurable $f : x \to \mathbb{C}$ for which $\|f\|_p < \infty$. A special case is when $\mathcal{M} = \mathcal{P}(X)$ and μ is counting measure, in which case we write $l^p(X)$ for $L^p(X, \mathcal{M}, \mu)$. When also $X = \mathbb{N}$, we write simply l^p. (I.e. when we write l^p without specifying X, it is understood that $X = \mathbb{N}$.)

Clearly $\|\lambda f\|_p = |\lambda|\|f\|_p$ for any scalar λ. Also

$$\|f + g\|_p^p = \int |f + g|^p \leq \int (2 \max(|f|, |g|))^p \leq 2^p(\|f\|_p^p + \|g\|_p^p).$$

Thus L^p is a vector space (with identification of functions that are equal a.e.). We claim that $\|\cdot\|_p$ is a norm. What remains to be shown is that the triangle inequality holds.

Lemma 2.1 *Let $a, b \geq 0$ and $\lambda \in (0, 1)$. Then*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1 - \lambda) b.$$

Proof. The function $g(x) = a^x b^{1-x}$ is convex, so $g(\lambda) \leq \lambda g(1) + (1 - \lambda) g(0)$. This is exactly what we wanted to prove. $\square$

Theorem 2.2 *(***Hölder's inequality***) Let* $1 < p < \infty$ *and* $1/p + 1/q = 1$*. Then*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Remark. Below, we will define $\|\cdot\|_\infty$ in such a way that it will be easily seen that Hölder's inequality extends to $1 \leq p \leq \infty$.

Remark. When $1/p + 1/q + 1$, one says that p and q are *conjugate exponents*.

Proof. If either $\|f\|_p$ or $\|g\|_q$ is 0 or ∞, then the result is trivial, so assume otherwise. Apply Lemma 2.1 with $a = |f(x)|^p / \|f\|_p^p$, $b = |g(x)|^q / \|g\|_q^q$ and $\lambda = 1/p$, so that $1 - \lambda = 1/q$, to get

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

Integrating both sides gives

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

□

We are ready for the triangle inequality for L^p-norm a.k.a. Minkowski's inequality

Theorem 2.3 *(***Minkowski's inequality***) Suppose* $1 \leq p < \infty$*. Then*

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

Proof. The result is trivial for $p = 1$ or $f + g = 0$ a.e. so assume $1 < p < \infty$ and $\|f + g\|_p > 0$. Clearly

$$|f + g|^p \leq |f + g|^{p-1}(|f| + |g|).$$

Let $q = p/(p-1)$, the conjugate exponent of p. Then

$$\|f + g\|_p^p = \int |f + g|^p \leq \int |f||f + g|^{p-1} + \int |g||f + g|^{p-1}$$

$$\leq \|f\|_p \||f + g|^{p-1}\|_q + \|g\|_p \||f + g|^{p-1}\|_q = (\|f\|_p + \|g\|_p)\Big(\int |f + g|^p\Big)^{(p-1)/p}$$

$$= (\|f\|_p + \|g\|_p)\|f + g\|_p^{p-1},$$

where the second inequality uses Hölder's inequality and the first equality uses that $q(1-p) = p$. The result now follows on dividing both sides with $\|f+g\|_p^{p-1}$. $\square$

By Minkowski's inequality $\|\cdot\|_p$ is indeed a norm and L^p is indeed a NVS. Moreover

Theorem 2.4 *L^p is a Banach space.*

Proof. We want to show that L^p is complete and to that end we use Theorem 1.1. Assume that $\{f_n\}$ is absolutely convergent, i.e. $B := \sum_n \|f_n\|_p < \infty$. Write $G_n = \sum_1^n |f_k|$ and $G = \lim_n G_n = \sum_1^\infty |f_n|$. By Minkowski's inequality, $\|G_n\|_p \leq \sum_1^n \|f_k\|_p \leq B$. By the MCT, $\|G\|_p \leq B$. Hence $G \in L^p$ and $\sum_1^\infty f_k$ converges a.e. Write $F = \sum_1^\infty f_k \in L^p$. It remains to show that $\|\sum_1^n f_k - F\|_p \to 0$. However $F - \sum_1^n f_k = \sum_{n+1}^\infty f_k \to 0$ a.e. and

$$|F - \sum_1^n f_k|^p \leq \Big(|F| + \sum_1^k |f_k|\Big)^p \leq (2G)^p \in L^p$$

so this follows from the DCT. $\square$

The next result is an approximation result for L^p.

Proposition 2.5 *The family of simple functions, i.e.*

$$\{\phi = \sum_1^n a_j \chi_{E_j} : n \in \mathbb{N}, \{E_j\} \textit{ disjoint}\}$$

is dense in L^p.

Proof. Fix an arbitrary $f \in L^p$. A simple function ϕ is in L^p if $\mu(E_j) = 0$ for every j and it is a fundamental fact of integration theory that one can find such ϕ_n so that $\phi_n \to f$ and $|\phi_n| \uparrow f$ a.e. Since $|\phi_n - f|^p \leq (2|f|)^p$, the DCT gives $\int |\phi_n - f|^p \to 0$ as desired. $\square$

The space L^∞. Using the usual convention $\inf \emptyset = \infty$, we define

$$\|f\|_\infty = \operatorname*{ess\,sup}_x |f(x)| = \inf\{a : \mu\{x : |f(x)| > a\} > 0\}.$$

Note that $\mu\{x : |f(x)| > \|f\|_\infty\} = 0$.

We now define $L^\infty(X, \mathcal{M}, \mu)$ as the space of measurable functions $f : X \to \mathbb{C}$ such that $\|f\|_\infty < \infty$. Note that by the definition of $\|\cdot\|_\infty$, for any $f \in L^\infty$, there exists a bounded function f' such that $f = f'$ a.e., so L^∞ can be regarded as the space of bounded functions.

A few simple facts follow.

(1) $\|fg\|_1 \leq \|f\|_1\|g\|_\infty$ (i.e. Hölder's inequality holds also for $p \in \{1, \infty\}$.)

(2) $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ (i.e. Minkowski's inequality holds also for $p = \infty$.)

(3) $\|f_n - f\|_\infty \to 0 \Rightarrow \exists E$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E.

(4) The family of simple functions is dense in L^∞.

(5) L^∞ is a Banach space.

Statements 1-4 are very easy to prove and are left as an exercise. For (5), suppose that $\{f_n\}$ is absolutely convergent. Then $\{f_n(x)\}$ is absolutely convergent for a.e. x, so there exists a limit $f := \lim_n f_n$ a.e. Pick m so large that $n \geq m \Rightarrow \|f_n - f_m\|_\infty < 1$ and M so large that $|f_m(x)| < M$ a.e. Then clearly $|f(x)| < M + 1$ a.e., i.e. $f \in L^\infty$. Also $\|f_n - f\|_\infty \leq \sum_{n+1}^\infty \|f\|_\infty \to 0$.

Next we turn to relationships between L^p's for different p's. When $p < q < r$, we have on one hand that $L^q \subseteq L^p + L^r$, which means that any function in L^q can be written as the sum of one function in L^p and one function in L^r. On the other hand we have that $L^q \supseteq L^r \cap L^p$. The proofs of these two facts are left as exercises. Here are two more facts.

Proposition 2.6 *Assume that* $1 \leq p < q \leq \infty$.

(a) If μ is counting measure, then $\|f\|_p \geq \|f\|_q$. Consequently $L^p(\mu) \subseteq L^q(\mu)$.

(b) If μ is finite, then

$$\|f\|_p \leq \|f\|_q \mu(X)^{(q-p)/(pq)}.$$

Consequently $L^p \supseteq L^q$.

Remark. Part (b) when μ is a probability measure should be known to a probabilistically oriented person: $\mathbb{E}[\xi^p]^{1/p} \leq \mathbb{E}[\xi^q]^{1/q}$.

Proof.

(a) In case $q = \infty$, we have
$$\|f\|_\infty^p = (\sup_x |f(x)|)^p \le \sum_x |f(x)|^p = \|f\|_p^p.$$
Now assume $q < \infty$ and assume without loss of generality that $\|f\|_q = 1$. Then $|f(x)| \le 1$ for every x, so that $|f(x)|^p \ge |f(x)|^q$ and hence
$$\|f\|_p = \Big(\sum_x |f(x)|^p\Big)^{1/p} \ge \Big(\sum_x |f(x)|^q\Big)^{1/p} = 1 = \|f\|_q.$$

(b) If $q = \infty$, scale f so that $\|f\|_\infty = 1$. Then obviously $\|f\|_p \le \mu(X)^{1/p}$ as desired. For $q < \infty$, observe that $\dfrac{1}{q/p} + \dfrac{1}{q/(q-p)} = 1$. Hence Hölder's inequality gives
$$\|f\|_p^p = \||f|^p \cdot 1\|_1 \le \||f|^p\|_{q/p}\|1\|_{q/(q-p)} = \|f\|_q^p \mu(X)^{(q-p)/q}.$$
Now take both sides to the power $1/p$ to finish the proof.

$\square$

Example. L^p**-contraction of Markov chains.** Let $\{X_t\}_{t=0}^\infty$ be an irreducible aperiodic Markov chain on the finite state space S. Let π denote the stationary distribution (i.e. $\mathbb{P}(X_t = s) \to \pi(s)$, $s \in S$.) Let P be the transition matrix and let $f : S \to \mathbb{C}$. Write $\mathbb{E}_s[\cdot]$ for $\mathbb{E}[\cdot|X = s]$. Then $(Pf)(s) = \mathbb{E}_s[f(X_1)]$. For the $L^p(S, \mathcal{P}(S), \pi)$-norm, this means that
$$\begin{aligned}\|Pf\|_p^p &= \mathbb{E}_\pi\Big[|(Pf)(X_0)|^p\Big] = \mathbb{E}_\pi\Big[|\mathbb{E}_{X_0}[f(X_1)]|^p\Big] \\ &\le \mathbb{E}_\pi\Big[\mathbb{E}_{X_0}[|f(X_1)|]\Big] = \mathbb{E}_\pi\Big[|f(X_0)|\Big] = \|f\|_p^p\end{aligned}$$
with equality iff f is constant. Hence $\|Pf\|_p/\|f\|_p \le 1$. Better estimates of this ratio are sometimes used to bound the mixing time of the Markov chain. $\square$

3 The dual of L^p

Let $\mathcal{X}$ be a NVS. A linear map $\phi : \mathcal{X} \to K$ is called a (K-valued) *linear functional*. Denote by $\mathcal{X}^*$ the space $L(\mathcal{X}, K)$, the space of bounded linear functionals on $\mathcal{X}$. The space $\mathcal{X}^*$ is called the *dual of* $\mathcal{X}$.

In this section we will consider the case $\mathcal{X} = L^p(X, \mathcal{M}, \mu)$. The number q will throughout be assumed to be the conjugate exponent of p. For $g \in L^q$, define

$$\phi_g(f) = \int fg \, d\mu, \; f \in L^p.$$

Clearly ϕ_g is linear and

$$|\phi_g(f)| = \Big| \int fg \Big| \leq \int |fg| = \|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Hence $\|\phi_g\| \leq \|g\|_q$, and in particular ϕ_g is bounded, i.e. $\phi_g \in (L^p)^*$.

Proposition 3.1 *If $p > 1$, then $\|\phi_g\| = \|g\|_q$.*

Proof. Since $p > 1$ we have $q < \infty$. By the above, it remains to show that $\|g\|_q \leq \|\phi_g\|$. This is trivial if $g = 0$ a.e., so we may assume that $\|g\|_q > 0$. Let

$$f = \frac{|g|^{q-1}\overline{\mathrm{sgn} g}}{\|g\|_q^{q-1}}.$$

(Here the $\mathrm{sgn}(z) = e^{i \arg z} \Leftrightarrow z = |z| \mathrm{sgn} z \Leftrightarrow |z| = z \, \overline{\mathrm{sgn} z}$.) Then, since $(q-1)p = q$,

$$\|f\|_p^p = \frac{1}{\|g\|_q^{(q-1)p}} \int |g|^{(q-1)p} = \frac{\|g\|_q^q}{\|g\|_q^q} = 1.$$

Hence

$$\|\phi_g\| \geq |\phi_g(f)| = \Big| \int fg \Big| = \frac{\int |g|^{q-1} g \, \overline{\mathrm{sgn} g}}{\|g\|_q^{q-1}} = \frac{\int |g|^q}{\|g\|_q^{q-1}} = \|g\|_q.$$

□

Extension: If μ is semifinite, then the result of Proposition 3.1 holds also for $p = 1 \Leftrightarrow q = \infty$. (Recall that μ is semifinite if $\exists E : 0 < \mu(E) < \infty$.) We omit the proof.

Put in other words, Proposition 3.1 says that the map $g \to \phi_g$ is an isometry from L^q to $(L^p)^*$. In fact, it is also an isomorphism:

Theorem 3.2 *Let μ be σ-finite and $1 \leq p < \infty$. Then for any $\phi \in (L^p)^*$, there exists $g \in L^q$ such that $\phi = \phi_g$.*

Extension: If $1 < p < \infty$, then the result holds without restriction on μ. We will not prove this.

Moral: $(L^p)^*$ and L^q are isometrically invariant. In other words these can be seen as two interpretations of the same space. In short this is written $(L^p)^* \cong L^q$, but one often, with abuse of notation, writes even $(L^p)^* = L^q$.

Proof. Assume first that μ is finite. Fix an arbitrary $\phi \in (L^p)^*$. Since μ is finite, $\chi_E \in L^p$ for all $E \in \mathcal{M}$. Hence we can define $\nu(E) = \phi(\chi_E)$, $E \in \mathcal{M}$. Now suppose that E_j, $j = 1, 2, \dots$ are disjoint and write $E = \cup_1^\infty E_j$. Then

$$\|\chi_E - \sum_1^n \chi_{E_j}\|_p = \|\sum_{n+1}^\infty E_j\|_p = \mu\Big(\bigcup_{n+1}^\infty E_j\Big)^p \to 0.$$

Hence $\sum_1^n \chi_{E_j} \to \chi(E)$ in L^p. Since ϕ is a continuous functional, this entails that $\phi(\sum_1^n \chi_{E_j}) \to \phi(\chi(E))$ in L^p. By the linearity of ϕ, it follows that $\nu(\cup_1^\infty E_j) = \sum_1^\infty \nu(E_j)$, i.e. ν is a signed measure. Hence the Radon-Nikodym Theorem implies that there exists $g \in L^1$ such that $\nu(E) = \phi(\chi_E) = \int_E g\, d\mu$, $E \in \mathcal{M}$. By linearity of ϕ it follows that $\phi(f) = \int fg\, d\mu$ for all simple functions f. For general f, Proposition 2.5 tells us that there exist simple functions f_n such that $\|f_n - f\|_p \to 0$. Hence, by continuity of ϕ,

$$\phi(f) = \lim_n \phi(f_n) = \lim_n \int f_n g.$$

Also, By Hölder,

$$\Big|\int f_n g - \int fg\Big| \le \int |(f_n - f)g| \le \|f_n - f\|_p \|g\|_q \to 0.$$

Thus $\lim_n \int f_n g = \int fg$ and hence $\phi(f) = \int fg$ as desired. (Note however that there is a gap here, since we have not shown that $g \in L^q$, merely that $g \in L^1$. We will come back to that immediately after this proof.)

Now let μ be σ-finite and pick sets E_n such that $E_n \uparrow X$ and $\mu(E_n) < \infty$. By the above there exist $g_n \in L^q$ such that $\phi(f) = \int fg_n$, $f \in L^p(E_n)$ and $\|g\|_q = \|\phi|_{L^p(E_n)}\| \le \|\phi\|$. Note that each g_n is unique up to alterations on a null set. Define g by letting $g = g_n$ on E_n. Then $|g_n| \uparrow |g|$ a.e. Hence

$$\int |g|^q = \lim_n |g_n|^q \le \|\phi\|^q$$

so that $g \in L^q$. By the DCT, $g_n \to g$ in L^q. Now for $f \in L^p$, we have $f\chi_{E_n} \in L^p(E_n)$ and by the DCT $f\chi_{E_n} \to f$ in L^p. Hence

$$\left|\phi(f) - \int fg\right| = \left|\lim_n \phi(f\chi_{E_n}) - \int fg\right| = \left|\lim_n \int fg_n - \int fg\right|$$

$$\leq \limsup_n \int |f(g_n - g)| \leq \limsup_n \|f\|_p \|g_n - g\|_q = 0,$$

by Hölder's inequality. □

Now we fill in the gap in the above proof. We do so by using the following result, which is sometimes referred to as the "Reverse of Hölder's inequality".

Theorem 3.3 *Assume that μ is semifinite. Let S be the family of simple functions whose support is of finite measure. Assume that $fg \in L_1$ for all $f \in S$ and that*

$$M_q(g) := \sup\left\{\left|\int fg\, d\mu\right| : f \in S, \|f\|_p = 1\right\} < \infty.$$

Then $\|g\|_q = M_q(g)$. In particular $g \in L^q$.

Proof. That $M_q(g) \leq \|g\|_q$ follows from Hölder, so we focus on the reverse inequality. We will settle for the case when μ is σ-finite. Assume first that $q < \infty$. Let $E_n \uparrow X$ with $\mu(E_n) < \infty$. Let h_n be simple functions with $h_n \to g$ and $|h_n| \leq |g|$. Let $g_n = h_n \chi_{E_n}$. Then $g_n \in S$, $g_n \to g$ and $|g_n| \leq |g|$. Let

$$f_n = \frac{|g_n|^{q-1}\overline{\operatorname{sgn} g}}{\|g_n\|_q^{q-1}}.$$

Then, as in the proof of Theorem 3.1, $\|f_n\|_p = 1$. Note also that $\int |f_n g_n| = \|g\|_q$. Thus, by Fatou's Lemma,

$$\|g\|_q \leq \liminf_n \|g_n\|_q = \liminf_n \int |f_n g_n|$$

$$\leq \liminf_n \int |f_n g| = \liminf_n \int f_n g \leq M_q(g)$$

since $f_n g$ is nonnegative.

Now to the case $q = \infty$. Let $A = \{x : |g(x)| > M_\infty(g) + \epsilon\}$ and assume that $\mu(A) > 0$ for some $\epsilon > 0$. Then we can find $B \subseteq A$ with $0 < \mu(B) < \infty$. Let $f = \chi_B \overline{\text{sgn} g}/\mu(B)$. Then $f \in S$ and $\|f\|_p = \|f\|_1 = 1$. However

$$\int fg = \int_B \frac{|g|}{\mu(B)} > M_\infty(g) + \epsilon,$$

a contradiction. □

Now consider the function g from the proof of Theorem 3.2. For $f \in S$ and $\|f\|_p = 1$, we have $|\int fg| = |\phi(f)| \leq \|\phi\|$, so $M_g(q) < \infty$ and hence $g \in L^q$.

4 The Hahn-Banach Theorem

Let $\mathcal{X}$ be a NVS and recall that $\mathcal{X}^* = L(\mathcal{X}, K)$. How can we be sure that this is an interesting space in the sense that it really contains any nontrivial objects? In case $\mathcal{X} = L^p(X, \mathcal{M}, \mu)$ for some concrete $\mathcal{X}$ like e.g. $\mathcal{X} = \mathbb{R}^n$ we know, and saw in the previous section, that this is indeed the case, but in general? As we shall see, the Hahn-Banach Theorem answers our question with a "yes". First however, we need some preliminaries.

Suppose that $f : \mathcal{X} \to \mathbb{C}$ is a linear functional. Then $u := \Re f$ is a real-valued linear functional and since $\Im f(x) = -\Re(if(x)) = -\Re f(ix) = -u(ix)$, we have

$$f(x) = u(x) - iu(ix).$$

On the other hand, if $u : \mathcal{X} \to \mathbb{C}$ is a real-valued linear functional, then $f(x) := u(x) - iu(ix)$, $x \in \mathcal{X}$, is a complex-valued linear functional. In any case, the equality $f(x) = u(x) - iu(ix)$, $x \in \mathcal{X}$, entails that

$$\|f\| = \|u\|.$$

This follows on one hand from $|u(x)| = |\Re f(x)| \leq |f(x)|$, so that $\|u\| \leq \|f\|$, and on the other hand from

$$|f(x)| = f(x)\overline{\text{sgn} f(x)} = f(\overline{\text{sgn} f(x)}x) = u(\overline{\text{sgn} f(x)}x) \leq \|u\|\|x\|$$

so that $\|f\| \leq \|u\|$.

Let $\mathcal{X}$ be a vector space over $\mathbb{R}$. A *Minkowski functional* on $\mathcal{X}$ is a function $p : \mathcal{X} \to \mathbb{R}$ such that

- $p(\lambda x) = \lambda p(x)$ for all $\lambda \geq 0$ and $x \in \mathcal{X}$,
- $p(x+y) \leq p(x) + p(y)$ for all $x, y \in \mathcal{X}$.

In particular all linear functionals and all seminorms are Minkowski functionals.

Theorem 4.1 *(***Hahn-Banach Theorem, real version***) Let $\mathcal{X}$ be a vector space over $\mathbb{R}$ and p a Minkowski functional on $\mathcal{X}$. Let $\mathcal{M}$ be a linear subspace of $\mathcal{X}$ and $f : \mathcal{M} \to \mathbb{R}$ a linear functional such that $f \leq p$ on $\mathcal{M}$. Then there exists a linear functional $F : \mathcal{X} \to \mathbb{R}$ such that $F = f$ on $\mathcal{M}$ and $F \leq p$.*

Proof. The result is trivial if $\mathcal{M} = \mathcal{X}$, so assume that there exists an $x \in \mathcal{X} \setminus \mathcal{M}$. For any $y_1, y_2 \in \mathcal{M}$,

$$f(y_1) + f(y_2) = f(y_1 + y_2) \leq p(y_1 + y_2) \leq p(y_1 - x) + p(y_2 + x),$$

i.e.

$$f(y_1) - p(y_1 - x) \leq f(y_2) - p(y_2 + x).$$

Since y_1 and y_2 were arbitrary, we get

$$\sup_{y \in \mathcal{M}} \Big(f(y) - p(y - x) \Big) \leq \inf_{y \in \mathcal{M}} \Big(f(y) - p(y + x) \Big).$$

Fix a number α between these two quantities. Define $g : \mathcal{M} + \mathbb{R}x \to \mathbb{R}$ by $g(y + \lambda x) = f(y) + \lambda \alpha$. Then g is linear and extends f and for $\lambda > 0$,

$$\begin{aligned} g(y + \lambda x) = \lambda\Big(f(y/\lambda) + \alpha\Big) &\leq \lambda\Big(f(y/\lambda) + p(x + y/\lambda) - f(y/\lambda)\Big) \\ &= p(y + \lambda x), \end{aligned}$$

where the inequality follows from lower bound in the definition of α. For $\lambda < 0$,

$$\begin{aligned} g(y + \lambda x) = |\lambda|\Big(f(y/|\lambda|) - \alpha\Big) &\leq |\lambda|\Big(f(y/|\lambda|) - f(y/|\lambda|) + p(y/|\lambda| - x)\Big) \\ &= p(y + \lambda x) \end{aligned}$$

where the inequality follows from the upper bound in the definition of α. Hence $g \leq p$.

Now let $\mathcal{F}$ be the family of linear extensions, g, of p with $g \leq p$. Partially order $\mathcal{F}$ with respect to inclusion of domain. Then every chain has an upper bound, namely the extension defined on the union of the domains of the elements of the chain. By Zorn's Lemma, $\mathcal{F}$ has a maximal element, F. The element F must be defined on the whole of $\mathcal{X}$ for if not, it could be extended by the above procedure. $\square$

Now assume that $f : \mathcal{M} \to \mathbb{C}$ is a linear functional and that $|f| \leq p$, where p is a seminorm, i.e. a Minkowski functional with the extra property that $p(\lambda x) = |\lambda| p(x)$ for all $\lambda \in \mathbb{C}$ and $x \in \mathcal{X}$. Then $u = \Re f$ is a real-valued linear functional and can hence be extended to $U : \mathcal{X} \to \mathbb{C}$ with $|U| \leq p$. Let $F(x) = U(x) - iU(ix)$. Then f extends f to the whole of $\mathcal{X}$ and

$$|F(x)| = F(x)\overline{\mathrm{sgn} F(x)} = F(\overline{\mathrm{sgn} F(x)}x) = U(\overline{\mathrm{sgn} F(x)}x)$$

$$\leq p(\overline{\mathrm{sgn} F(x)}x) = p(x).$$

In summary:

Theorem 4.2 *(***Hahn-Banach Theorem, complex version***) Let $f : \mathcal{M} \to \mathbb{C}$ be a linear functional on the linear subspace $\mathcal{M}$ of $\mathcal{X}$. Assume that p is a seminorm on $\mathcal{X}$ and that $|f| \leq p$ on $\mathcal{M}$. Then there exists $F : \mathcal{X} \to \mathbb{C}$ such that $F = f$ on $\mathcal{M}$ and $|F| \leq p$.*

Here are some applications.

(a) For any $x \in \mathcal{X}$ there exists an $f \in \mathcal{X}^*$ such that $f(x) = \|x\|$ and $\|f\| = 1$.

Proof. Let $\mathcal{M} = \mathbb{C}x$ and define f on $\mathcal{M}$ by $f(\lambda x) = \lambda\|x\|$. Define p by $p(y) = \|y\|$. Then p is a norm andS hence a seminorm and $|f(\lambda x)| = |\lambda\|x\|| = \|\lambda x\| = p(x)$. Now extend f to the whole of $\mathcal{X}$ using the Hahn-banach Theorem. The extension satisfies $|f| \leq p$, so for any $y \in \mathcal{X}$, $|f(y)| \leq \|y\|$, so $\|f\| \leq 1$. However $|f(x)| = \|x\|$, so $\|f\| = 1$. $\square$

(b) If $x_1, x_2 \in \mathcal{X}$ and $x_1 \neq x_2$, then there exists $f \in \mathcal{X}^*$ such that $f(x_1) \neq f(x_2)$.

Proof. Use (a) with $x = x_1 - x_2$. $\square$

(c) If $\mathcal{M}$ is a closed subspace of $\mathcal{X}$ and $x \in \mathcal{X} \setminus \mathcal{M}$, then there exists $f \in \mathcal{X}^*$ with $f \equiv 0$ on $\mathcal{M}$ and $f(x) \neq 0$.

Proof. Let $\delta = \inf\{\|x-y\| : y \in \mathcal{M}\}$. Then $\delta > 0$, for if $\delta = 0$, then there would exist $y_n \in \mathcal{M}$ with $\|x - y_n\| \to 0$, i.e. $y_n \to x$. Since $\mathcal{M}$ is closed this entails that $x \in \mathcal{M}$, a contradiction.

Define f on $\mathcal{M} + \mathbb{C}x$ by

$$f(y + \lambda x) = \lambda\delta.$$

Then $f \equiv 0$ on $\mathcal{M}$ and

$$|f(y + \lambda x)| = |\lambda|\delta \leq |\lambda|\|x + y/\lambda\| = \|y + \lambda x\|,$$

where the inequality follows from the definition of δ. Thus $|f| \leq \|\cdot\|$, so f can be extended by Hahn-Banach. $\square$

(d) For $x \in \mathcal{X}$, define $\hat{x} : \mathcal{X}^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$. Then $x \to \hat{x}$ is a linear isometry from $\mathcal{X}$ to $\mathcal{X}^{**}$. In particular, letting $\hat{\mathcal{X}} := \{\hat{x} : x \in \mathcal{X}\}$, we have $\hat{\mathcal{X}} \subseteq \mathcal{X}^{**}$. A NVS $\mathcal{X}$ for which $\hat{\mathcal{X}} = \mathcal{X}^{**}$ is said to be *reflexive*.

Proof. Clearly the given map is linear. Also

$$|\hat{x}(f)| = |f(x)| \leq \|f\|\|x\|$$

and hence $\|\hat{x}\| \leq \|x\|$. On the other hand, by (a) there exists an $f \in \mathcal{X}^*$ such that $f(x) = \|x\|$ and $\|f\| = 1$, so $|\hat{x}(f)| = \|x\| = \|x\|$ and hence $\|\hat{x}\| \geq \|x\|$. $\square$

5 The Baire Category Theorem and consequences thereof

The Baire Category Theorem (BCT) is the following topological result.

Theorem 5.1 *(**Baire Category Theorem**) Let X be a complete metric topological space. Then the following hold.*

(a) If for each $n = 1, 2, \ldots$, U_n is an open dense set, then $\cap_n U_n$ is dense.

(b) The space X cannot be written as a countable union of nowhere dense sets.

Proof. First we show that (a) implies (b). If E_n is nowhere dense, i.e. $\overline{E_n}^o = \emptyset$, then $\overline{E_n}^c$ is dense an open, so by (a) $\cap_n \overline{E_n}^c$ is dense. Therefore

$$\Big(\cup_n E_n\Big)^c = \cap_n E_n^c \supseteq \cap_n \overline{E_n}^c$$

and since a dense set cannot be empty, (b) follows.

Now we prove (a). We want to show that $(\cap_n U_n) \cap W \neq \emptyset$ for any open nonempty set W. Since $U_1 \cap W$ is open and nonempty, we can find $r_1 \in (0,1)$ and $x_1 \in X$ such that $\overline{B_{r_1}(x_1)} \subseteq U_1 \cap W$. In the same way there exist $r_2 \in (0,1/2)$ and $x_2 \in X$ such that $\overline{B_{r_2}(x_2)} \subseteq U_2 \cap B_{r_1}(x_1)$. Repeating the argument once again gives $r_3 \in (0,1/4)$ and $x_3 \in X$ such that $\overline{B_{r_3}(x_3)} \subseteq U_3 \cap B_{r_2}(x_2)$. Repeating the argument inductively gives a Cauchy sequence of points x_n. Since X is complete, $x_n \to x$ for some $x \in X$. For every n, we have

$$x \in \overline{B_{r_n}(x_n)} \subseteq U_n \cap B_{r_{n-1}}(x_{n-1}) \subseteq U_n \cap W.$$

Hence $x \in (\cap_n U_n) \cap W$. $\square$

One should note that the BCT is purely topological, it holds for any topological space that is homeomorphic to a complete metric space.

Example. Wiener's Tauberian Theorem. Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z} = \mathbb{R} \bmod 2\pi$ be the unit circle in the complex plane.

Question: For which functions $f \in L^1(\mathbb{T})$ is it true that the translates $f_y = f(\cdot - y)$, $y \in \mathbb{T}$, span $L^1(T)$ in the sense that the set

$$M := \{\sum_{j=1}^{N} a_j f_{y_j} : N \in \mathbb{N}, a_j \in \mathbb{C}, y_j \in \mathbb{T}\}$$

is dense in $L^1(\mathbb{T})$?

Let $c_k = \int_{\mathbb{T}} f(x)e^{-ikx}dx$, f's k'th Fourier coefficient. Note that the k'th Fourier coefficient of f_y is $e^{-iky}c_k$. If $c_k = 0$ for some k, this means that the k'th Fourier coefficient of f_y is also 0 for all $y \in \mathbb{T}$ and hence that all functions in M have 0 for their k't Fourier coefficient. Hence e.g. the function $e^{ikx} \in L^1(T)$, whose k't Fourier coefficient is 1, cannot be approximated by functions in M.

Now assume $c_k \neq 0$ for all k. Assume that $\overline{M} \neq L^1(T)$ and pick $g_0 \in L^1(T) \setminus M$. According to application (c) above of Hahn-Banach, there exists a linear functional ϕ on $L^1(T)$ such that $\phi \equiv 0$ on $\overline{M}$ and $\phi(g_0) \neq 0$. By Theorem 3.2 there exists a function $h \in L^\infty(\mathbb{T})$ such that

$$\phi(g) = \int_{\mathbb{T}} g(x)h(x)dx, \; g \in L^1(\mathbb{T}).$$

Since $f_y \in M$ for all $y \in \mathbb{T}$, this entails that $\int_{\mathbb{T}} f(x-y)h(x)dx = 0$ for all $y \in \mathbb{T}$. In other words, the convolution $\check{h} * f$ is identically 0 (where $\check{h}(x) = h(-x)$). However the k'th Fourier coefficient of $\check{h} * f$ is $\overline{d_k} c_k$, where d_k is the k'th Fourier coefficient of h. Since $c_k \neq 0$ for all k, we have $d_k = 0$ for all k. Thus $h \equiv 0$ and consequently $\phi(g_0) = 0$, a contradiction.

In summary: M spans $L^1(\mathbb{T})$ iff $c_k \neq 0$ for all k. □

Recall that a map $f : X \to Y$ is called an *open map* if $f(U)$ is open in Y whenever U is open in X. An immediate consequence is that if f is bijective and open, then f^{-1} is continuous. If Y is metric, then openness of f can be expressed as that for all open U in X and all $x \in U$, there exists $r > 0$ such that

$$B_r(f(x)) \subseteq f(U).$$

If If X and Y are NVS's and f is linear, then this boils down to

$$\exists r > 0 : B_r(0) \subseteq f(B_1(0)).$$

Theorem 5.2 *(***The Open Mapping Theorem***) Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces and $T \in L(\mathcal{X}, \mathcal{Y})$. Then T surjective implies T open.*

Proof. Write B_r for $B_r(0)$ (in $\mathcal{X}$ as well as in $\mathcal{Y}$). We want to show that $\exists r > 0 : B_r \subseteq T(B_1)$. Since T is surjective and $\mathcal{X} = \cup_n B_n$, $\mathcal{Y} = \cup_n T(B_n) = \cup_n n T(B_1)$. Hence the BCT implies that $T(B_1)$ is not nowhere dense. (Here we use that $\mathcal{Y}$ is complete.) Thus we can find $y_0 \in \mathcal{Y}$ and $r > 0$ such that

$$B_{8r}(y_0) \subseteq \overline{T(B_1)}.$$

Pick $y_1 \in \overline{T(B_1)}$ such that $\|y_1 - y_0\| < 4r$ and $y_1 = Tx$ for some $x \in B_1$. Then $B_{4r}(y_1) \subseteq \overline{T(B_1)}$ and whenever $\|y\| < 4r$, we have $y = y_1 + (y - y_1) \in \overline{T(B_2)}$, since $y - y_1 \in B_{4r}(-y_1) \subseteq \overline{T(B_1)}$. By scaling, this generalizes to $y \in \overline{T(B_{2^{-n}})}$ whenever $\|y\| < 2r2^{-n}$.

Now pick y with $\|y\| < r$. Then $y \in \overline{T(B_{1/2})}$. Hence there exists $x_1 \in B_{1/2}$ with $\|y - Tx_1\| < r/2$. This entails that $y - Tx_1 \in \overline{T(B_{1/4})}$. Thus we can find

$x_2 \in B_{1/4}$ with $\|y - Tx_1 - Tx_2\| < r/4$. Then $y - Tx_1 - Tx_2 \in \overline{T(B_{1/8})}$, etc. Keeping on inductively gives $x_j \in B_{2^{-j}}$. Since $\mathcal{X}$ is Banach, $\sum_j x_j$ is convergent by Theorem 1.1. Let $x = \sum_j x_j$. Then $x \in B_1$ and $\|y - Tx\| = \lim_N \|y - \sum_1^N Tx_j\| = 0$, i.e. $y = Tx$. In summary there exists $x \in B_1$ such that $y = Tx$ whenever $\|y\| < r$, as desired. $\square$

Corollary 5.3 *If $\mathcal{X}$ and $\mathcal{Y}$ are Banach and $T \in L(\mathcal{X}, \mathcal{Y})$ is bijective, then T is an isomorphism.*

Proof. Since T is bijective, then T^{-1} exists. By the Open Mapping Theorem, T^{-1} is continuous and hence bounded. $\square$

Theorem 5.4 *(**The Closed Graph Theorem**) Assume that $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, $T : \mathcal{X} \to \mathcal{Y}$ linear and that*

$$\Gamma(T) := \{(x, Tx) : x \in \mathcal{X}\} \subseteq \mathcal{X} \times \mathcal{Y}$$

is closed. Then T is bounded.

Proof. Let π_1 and π_2 be the projection maps from $\Gamma(T)$ to $\mathcal{X}$ and $\mathcal{Y}$ respectively, given by $\pi_1(x, Tx) = x$ and $\pi_2(x, Tx) = Tx$. By the definition of product norm, $\|\pi_1\| \leq 1$ and $\|\pi_2\| \leq 1$ and hence $\pi_1 \in L(\Gamma(T), \mathcal{X})$ and $\pi_2 \in L(\Gamma(T), \mathcal{Y})$.

Note that $\Gamma(T)$ is complete, since it is a closed subspace of the complete space $\mathcal{X} \times \mathcal{Y}$. Thus, since π_1 is bijective, it follows that π_1^{-1} is bounded by the above corollary. Since also π_2 is bounded, $\pi_2 \circ \pi_1^{-1} = T$ is bounded. $\square$

Theorem 5.5 *(**The Uniform Boundedness Principle**) Assume that $\mathcal{X}$ is a Banach space and $\mathcal{Y}$ a NVS. Let $\mathcal{A}$ be a subfamily of $L(\mathcal{X}, \mathcal{Y})$. If $\sup_{T \in \mathcal{A}} \|Tx\| < \infty$ for all $x \in \mathcal{X}$, then $\sup_{T \in \mathcal{A}} \|T\| < \infty$.*

Proof. Let

$$E_n = \{x : \sup_{T \in \mathcal{A}} \|Tx\| \leq n\} = \bigcap_{T \in \mathcal{A}} \{x : \|Tx\| \leq n\}.$$

Observe that by assumption $\cup_n E_n = \mathcal{X}$ and that since $\|T(\cdot)\|$ is continuous, each E_n is closed. By the BCT, some E_n must contain a ball, $\overline{B_r(x_0)}$. For any x with $\|x\| \leq r$, we have for any $T \in \mathcal{A}$,

$$\|Tx\| \leq \|T(x - x_0)\| + \|Tx_0\| \leq 2n,$$

since $x - x_0$ and x_0 are both in $\overline{B_r(x_0)} \subseteq E_n$. Hence for any x with $\|x\| \leq 1$ and any $T \in \mathcal{A}$, we have $\|Tx\| \leq 2n/r$. Hence $\sup_{T\in\mathcal{A}} \|T\| \leq 2n/r$. $\square$

6 Hilbert spaces

Hilbert spaces are Banach spaces with extra geometric structure. The norm arises from an *inner product*. Let $\mathcal{X}$ be a vector space over $\mathbb{C}$. An operation $\langle\cdot,\cdot\rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is called an inner product if for all $x, y, z \in \mathcal{X}$ and $a, b \in \mathbb{C}$,

(a) $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$,

(b) $\langle y, x\rangle = \overline{\langle y, x\rangle}$,

(c) $\langle x, x\rangle \in (0, \infty)$ for all $x \neq 0$.

Note that from (a) and (b), $\langle x, ay + bz\rangle = \overline{\langle ay + bz, x\rangle} = \overline{a\langle y, x\rangle + b\langle z, x\rangle} = \overline{a}\langle x, y\rangle + \overline{b}\langle x, z\rangle$. It also follows that $\langle 0, x\rangle = \langle x, 0\rangle = 0$ for all x. When $\mathcal{X}$ is equipped with an inner product, we say that $\mathcal{X}$ is a *pre-Hilbert space*.

Assume now that $\mathcal{X}$ is a pre-Hilbert space. Define

$$\|x\| = \langle x, x\rangle^{1/2}.$$

The notation suggests that this is a norm. Let us show that this is indeed the case. For that we will need the following well-known result.

Theorem 6.1 *(***Schwartz' inequality***) For all $x, y \in \mathcal{X}$,*

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

with equality iff x and y are linearly dependent.

Proof. The result is trivial if $y = 0$, so we may assume that $y \neq 0$. Let $\alpha = \operatorname{sgn}\langle x, y\rangle$, $z = \alpha y$ and pick $t \in \mathbb{R}$. Then

$$0 \leq \langle x - tz, x - tz\rangle = \langle x, x\rangle - t\langle x, z\rangle - t\langle z, x\rangle + t^2\langle z, z\rangle$$

$$= \|x\|^2 - 2t|\langle x, y\rangle| + t^2\|y\|^2$$

where the last equality follows from the definition of α. Differentiate the right hand side with respect to t and derive that this expression is minimized for $t = |\langle x, y\rangle|/\|y\|^2$. Plugging this into the inequality yields,

$$0 \leq \|x\|^2 - \frac{2|\langle x, y\rangle|^2}{\|y\|^2} + \frac{|\langle x, y\rangle|^2}{\|y\|^2} = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}.$$

This proves the desired inequality. Also, by part (c) in the definition of inner product, equality holds iff $x - tz = x - \alpha t y = 0$ as desired. □

Note. If $\langle \cdot, \cdot \rangle$ is real-valued, then the t defined in the proof is also real-valued. Hence equality holds in Schwarz' inequality iff $x = cy$ for a *real* constant c.

Note that $\langle x, y\rangle + \langle y, x\rangle = 2\Re\langle x, y\rangle$. Therefore

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

$$\leq \|x\|^2 + 2|\langle x, y\rangle|\|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

here the second inequality follows from Schwarz' inequality. This proves the triangle inequality. Hence $\|\cdot\|$ is a norm. Thus a pre-Hilbert space $\mathcal{X}$ is a NVS. If $\mathcal{X}$ is also Banach, then $\mathcal{X}$ is said to be a *Hilbert space*.

Example. Consider $\mathcal{X} = L^2(X, \mathcal{M}, \mu)$. Let

$$\langle f, g\rangle = \int f\overline{g}\, d\mu.$$

This is well defined since $|f\overline{g}| \leq (|f|^2 + |g|^2)/2$. It is readily checked that this defines an inner product, and the resulting norm becomes

$$\|f\| = (\int f\overline{f})^{1/2} = (\int |f|^2)^{1/2} = \|f\|_2.$$

Hence L^2 can be seen as a Hilbert space. (However L^p is not a Hilbert space for any p other than 2.) □

Proposition 6.2 *$\langle \cdot, \cdot \rangle$ is continuous.*

Proof. Let $x_n \to x$ and $y_n \to y$, i.e. $\|x_n - x\| \to 0$ and $\|y_n - y\| \to 0$. Then

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| = |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \leq \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\| \to 0,$$

since $\{\|y_n\|\}$ converges to $\|y\|$ and is hence bounded. $\square$

Whenever x and y are two elements of the Hilbert space $\mathcal{X}$ such that $\langle x, y\rangle = 0$, we say that x and y are *orthogonal*. For short, this is sometimes written as $x \perp y$. The following version of the Pythagorean Theorem holds for Hilbert spaces. The proof is trivial.

Theorem 6.3 *(***The Pythagorean Theorem***) If $x_i \perp x_j$ for all $1 \leq i < j \leq n$, then*

$$\|\sum_1^n x_j\|^2 = \sum_1^n \|x_j\|^2.$$

We also have

Theorem 6.4 *(***The Parallelogram Law***) For all $x, y \in \mathcal{X}$,*

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

The proof is a straightforward expansion of the left-hand side.

The next result states that vectors in a Hilbert space can be projected onto a closed linear subspace. For a linear subspace $\mathcal{M}$ of $\mathcal{X}$, we write $\mathcal{M}^\perp := \{x \in \mathcal{X} : \forall y \in \mathcal{M} : x \perp y\}$.

Theorem 6.5 *Let $\mathcal{M} \subseteq \mathcal{X}$ be a closed linear subspace. Then for any $x \in \mathcal{X}$ there exist unique elements $y \in \mathcal{M}$ and $z \in \mathcal{M}^\perp$ such that $x = z + y$.*

Proof. Fix x and let $\delta = \inf\{\|x - y\| : y \in \mathcal{M}\}$. Let $y_n \in \mathcal{M}$ be such that $\|x - y_n\| \to \delta$. By the Parallelogram Law,

$$2\|y_n - x\|^2 + 2\|y_m - x\|^2 = \|y_n - y_m\|^2 + \|y_n + y_m - 2x\|^2.$$

Rearranging gives

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\|\frac{1}{2}(y_n + y_m) - x\|^2 \\ &\leq 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4\delta^2\end{aligned}$$

by the definition of δ, since $\frac{1}{2}(y_n + y_m) \in \mathcal{M}$. The right-hand side tends to 0 as $m, n \to \infty$, so $\{y_n\}$ is Cauchy and hence convergent. Let $y = \lim_n y_n$. Since $\mathcal{M}$ is closed, $y \in \mathcal{M}$. Let $z = x - y$. We claim that $z \in \mathcal{M}^\perp$. To see that this is so, pick $u \in \mathcal{M}$. Then there exists $c \in \mathbb{C}$ with $|c| = 1$ so that $\langle z, cu \rangle$ is real. Since $z + tcu = x - (y - tcu)$ and $y - tcu \in \mathcal{M}$, the function $f(t) := \|z + tcu\|^2$ is minimized for $t = 0$. However

$$f(t) = \|z\|^2 + 2t\langle z, cu \rangle + t^2\|u\|^2$$

so that $0 = f'(0) = 2\langle z, cu \rangle$. Hence also $\langle z, u \rangle = 0$ as desired.

For the uniqueness part, assume that we also have $x = z' + y'$. Then $y - y' = z' - z$ so that $y - y', z' - z \in \mathcal{M} \cap \mathcal{M}^\perp$, i.e. $y = y', z = z'$. □

Fixing an element $y \in \mathcal{X}$, we can define the linear functional $f_y(x) = \langle x, y \rangle$, $x \in \mathcal{X}$. Then, by Schwarz, $|f_y(x)| = |\langle x, y \rangle \leq \|x\|\|y\|$ with equality when x is a constant times y. Hence $\|f_y\| = \|y\|$ and the map $y \to f_y$ is a conjugate linear isometry from $\mathcal{X}$ to $\mathcal{X}^*$. In fact, the following result shows that it is also surjective.

Theorem 6.6 *For every $f \in \mathcal{X}^*$ there exists a unique $y \in \mathcal{X}$ such that $f(x) = \langle x, y \rangle$, $x \in \mathcal{X}$.*

Proof. If $\langle x, y \rangle = \langle x, y' \rangle$ for all $x \in \mathcal{X}$, then

$$\|y - y'\|^2 = \langle y - y', y \rangle - \langle y - y', y' \rangle = 0,$$

which proves uniqueness.

Now focus on the existence part. The case $f \equiv 0$ it trivial, so assume that $f \not\equiv 0$. Let $\mathcal{M} = \{x : f(x) = 0\}$. Then $\mathcal{M}$ is a closed linear subspace of $\mathcal{X}$. Hence, by Theorem 6.5, there exists $z \in \mathcal{M}^\perp$ such that $f(z) \neq 0$ and $\|z\| = 1$. Now pick an arbitrary $x \in \mathcal{X}$ and let $u = f(x)z - f(z)x$. Then $f(u) = 0$, so $u \in \mathcal{M}$ and hence $\langle u, z \rangle = 0$. In other words

$$0 = \langle f(x)z - f(z)x, z \rangle = f(x)\|z\|^2 - f(z)\langle x, z \rangle.$$

Solving for $f(x)$, recalling that $\|z\| = 1$, gives

$$f(x) = f(z)\langle x, z \rangle = \langle x, \overline{f(z)}z \rangle.$$

Taking $y = \overline{f(z)}z$ finishes the proof. □

So all Hilbert spaces $\mathcal{X}$ are isomorphic to $\mathcal{X}^*$ via the conjugate linear isometry $x \to f_y$. Hence $\mathcal{X} \cong \mathcal{X}^* \cong X^{**} \cong \ldots$. Moreover, recall the definition $\hat{x}(f) =$

$f(x)$ for a given $x \in \mathcal{X}$ and $\hat{\mathcal{X}} = \{\hat{x} : x \in \mathcal{X}\} \subseteq \mathcal{X}^{**}$. If $\mathcal{X}$ is Hilbert, then $\hat{x}(f) = \hat{x}(f_y) = \langle x, y\rangle$, i.e. $\hat{x} = \langle x, \cdot\rangle$. This entails that $\hat{\mathcal{X}} = \mathcal{X}^{**}$, i.e. $\mathcal{X}$ is reflexive.

The family $\{u_\alpha\}_{\alpha\in A}$ is said to an *orthonormal* family if $\langle u_{\alpha_1}, u_{\alpha_2}\rangle = 0$ for all $\alpha_1 \neq \alpha_2$ and $\|u_\alpha\| = 1$ for all α.

Theorem 6.7 *(***Bessel's inequality***) Suppose that $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal family. Then for all $x \in \mathcal{X}$,*

$$\sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2 \leq \|x\|^2.$$

In particular $\{\alpha : \langle x, u_\alpha\rangle \neq 0\}$ is countable.

Proof. Write $c_\alpha = \langle x, u_\alpha\rangle$. We have, for any finite $F \subseteq A$.

$$0 \leq \|x - \sum_{\alpha\in F} c_\alpha u_\alpha\|^2 = \|x\|^2 - 2\Re\langle x, \sum_{\alpha\in F} c_\alpha u_\alpha\rangle + \|\sum_{\alpha\in F} c_\alpha u_\alpha\|^2$$

$$\|x\|^2 - 2\sum_{\alpha\in F}|c_\alpha|^2 + \sum_{\alpha\in F}|c_\alpha|^2 = \|x\|^2 - \sum_{\alpha\in F}|c_\alpha|^2$$

where the equality follows from the Pythagorean Theorem. □

Theorem 6.8 *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal family. The following three statements are equivalent*

(a) $\forall\alpha : \langle x, u_\alpha\rangle = 0 \Rightarrow x = 0$ *(completeness),*

(b) $\forall x : \|x\|^2 = \sum_\alpha |\langle x, u_\alpha\rangle|^2$ *(Parseval's identity),*

(c) $\forall x : x = \sum_\alpha \langle x, u_\alpha\rangle u_\alpha$*, where the terms are nonzero for at most countably many* α*.*

Proof. That (b) implies (a) is trivial. To see that (a) implies (c), pick x and let $\alpha_1, \alpha_2, \ldots$ be an enumeration of the α's for which $\langle x, u_\alpha\rangle \neq 0$. Write $u_j = u_{\alpha_j}$. By Bessel's inequality, $\sum_j |\langle x, u_j\rangle|^2$ converges, so by the Pythagorean Theorem,

$$\|\sum_m^n \langle x, u_j\rangle u_j\|^2 = \sum_m^n |\langle x, u_j\rangle|^2 \to 0$$

as $m, n \to \infty$. Since $\mathcal{X}$ is complete, this means that $\sum_j \langle x, u_j \rangle u_j$ converges. Since $\langle x - \sum_j \langle x, u_j \rangle, u_\alpha \rangle = 0$ for all α, it follows from (a) that $x = \sum_j \langle x, u_j \rangle u_j$ as desired.

To see that (c) implies (b), by the same calculations as in the proof of Bessel,

$$\|x\|^2 - \sum_1^n |\langle x, u_j \rangle|^2 = \|x - \sum_1^n \langle x, u_j \rangle u_j\|^2 \to 0$$

as $n \to \infty$ by (c). $\square$

A family that satisfies the three statements in the above theorem, is called an orthonormal *basis* for $\mathcal{X}$.

Theorem 6.9 *Every Hilbert space has an orthonormal basis.*

Proof. Order the set of orthonormal families according to inclusion. Then every chain has a supremum, namely the union of the families in the chain. By Zorn's Lemma, there is a maximal element, $\{u_\alpha\}$. If this family does not span the whole of $\mathcal{X}$, there exists a $z \in \mathcal{X}$ such that $x := z - \sum_\alpha \langle z, u_\alpha \rangle u_\alpha \neq 0$. However, then $x/\|x\|$ can be added to the $\{u_\alpha\}$, contradicting maximality. $\square$

Theorem 6.10 *A Hilbert space has a countable basis iff it is separable.*

Remark. This result does not contradict exercise 26, since the notion of a basis is different there.

Proof. For the backwards implication, let $\{x_j\}$ be countable and dense. Let $y_1 = x_1$ and recursively $y_j = x_{n_j}$, where n_j is the first index k after n_{j-1} such that x_k is not in the span of $y_1, \ldots, y_{j-1}$. Then $\{y_j\}$ is linearly independent and dense. Now use the Gram-Schmidt process on the y_j's.

For the forwards direction, suppose that $\{u_j\}$ is a countable basis. Then it is easily seen that $\{\sum_1^N a_k u_{j_k} : n \in \mathbb{N}, a_k \in \mathbb{Q} + i\mathbb{Q}\}$ is dense. $\square$

Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal basis for the Hilbert space $\mathcal{X}$. Then the map $x \to \{\langle x, u_\alpha \rangle\}_{\alpha \in A}$ is an isometry from $\mathcal{X}$ to $l^2(A)$, by Parseval's identity. One can also show that this map is surjective. Hence $\mathcal{X} \cong l^2(A)$.

7 Weak and weak* convergence

Let $\mathcal{X}$ be a NVS and $\mathcal{X}^* = L(\mathcal{X}, \mathbb{C})$. The *weak toplogy* on $\mathcal{X}$ is generated by $\mathcal{X}^*$, i.e. $x_n \to x$ weakly iff $f(x_n) \to f(x)$ for every $f \in \mathcal{X}^*$.

The *weak* topology* on $\mathcal{X}^*$ is generated by $\hat{\mathcal{X}}$, i.e. $f_n \to f$ weakly* iff $f_n(x) \to f(x)$ for every $x \in \mathcal{X}$, i.e. if $f_n \to f$ pointwise.

Remark. When speaking of weak* convergence on a NVS $\mathcal{Y}$, one must always be able to see $\mathcal{Y}$ as the dual of some NVS $\mathcal{X}$, i.e. $\mathcal{Y} \cong \mathcal{X}^*$. One says that $\mathcal{X}$ is then a *pre-dual* of $\mathcal{Y}$.

Remark. Since $\hat{\mathcal{X}} \subseteq \mathcal{X}^{**}$, weak* convergence on $\mathcal{X}^*$ is weaker than weak convergence on $\mathcal{X}^*$. The two notions coincide when $\mathcal{X}$ is reflexive.

Example. Let $1 < p < \infty$. We may then consider weak and weak* convergence on L^p. Saying that $f_n \to f$ weakly in L^p means that $\phi(f_n) \to \phi(f)$ for every $\phi \in (L^p)^*$, i.e.

$$\int f_n g \to \int fg$$

for every $g \in L^q$, where q is the conjugate exponent of p. Saying that $f_n \to f$ weakly* on L^p means that we identify L^p with the dual of L^q. Hence $f_n \to f$ weakly* means $\phi_{f_n}(g) \to \phi_f(g)$, i.e.

$$\int f_n g \to \int fg$$

for every $g \in L^q$. We see that for $1 < p < \infty$, weak and weak* convergence on L^p agree.

In fact, this can be seen more directly by noting that L^p is reflexive: $(L^p)^{**} \cong (L^q)^* \cong L^p$.

For $p = 1$, weak convergence means $\int f_n g \to \int fg$ for every $g \in L^\infty$. However, weak* convergence cannot be similarly characterized, since $(L^\infty)^* \ncong L^1$. For l^1, we have $l^1 \cong c_0^* \cong c_{00}^*$ where $c_0 = \{g : \mathbb{N} \to \mathbb{C} : \lim_k g(k) = 0\}$ and $c_{00} = \{g : \mathbb{N} \to \mathbb{C} : g(k) = 0 \text{ eventually}\}$.

For $p = \infty$, weak* convergence means $\int f_n g \to \int fg$ for all $g \in L^1$. For weak convergence, no simple characterization can be made. $\square$

Theorem 7.1 **(Alaoglu's Theorem)** *Let $\mathcal{X}$ be a NVS and $f_n \in \mathcal{X}^*$ with $\|f_n\| \leq 1$, $n = 1, 2, \ldots$. Then there exists a convergent subsequence $\{f_{n_k}\}$.*

Remark. Since $\mathcal{X}^*$ is a complete metric space, sequential compactness and compactness are equivalent. Hence Alaoglu's Theorem can be stated as that the closed unit ball is compact in the weak* topology.

Proof. We will settle for the case when $\mathcal{X}$ is separable. Let $\{x_k\} \subseteq \mathcal{X}$ be dense. We want to show that there exists an $f \in \mathcal{X}^*$ such that f_{n_k} converges to f pointwise. Since $|f(x_1)| \le \|x_1\|$, there exists a subsequence $\{n_j^1\}$ such that $\{f_{n_j^1}(x_1)\}$ converges. Since $|f(x_2)| \le \|x_2\|$, there exists a further subsequence $\{n_j^2\}$ such that $\{f_{n_j^2}(x_2)\}$ also converges. Continuing this procedure indefinitely produces subsequences $\{n_j^i\}$ such that for each i, $\{f_{n_j^i}(x_k)\}$ converges for all $k \le i$. Now use diagonalization: $\{f_{n_j^j}(x_k)\}$ converges for every k. Let thus $n_k = n_j^j$. For $x \notin \{x_k\}$, fix $\epsilon > 0$ and pick x_k so that $\|x - x_k\| < \epsilon/3$. We have

$$|f_{n_i}(x) - f_{n_j}(x)| \le |f_{n_i}(x) - f_{n_i}(x_k)| + |f_{n_i}(x_k) - f_{n_j}(x_k)| + |f_{n_j}(x_k) - f_{n_j}(x)| < \epsilon$$

for large enough i and j, since the first and the third term are bounded by $\epsilon/3$, since $\|f_{n_i}\|$ and $\|f_{n_j}\|$ are bounded by 1. Hence $\{f_{n_k}(x)\}$ is Cauchy and hence convergent. Now set $f(x) = \lim_k f_{n_k}(x)$, $x \in \mathcal{X}$. $\square$

Example. The Poisson kernel in the unit disc. Let $U = \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disc in the complex plane and $\mathbb{T} = \partial U$. Recall that $f : \Omega \to \mathbb{C}$ is *harmonic* in Ω if $f''_{xx} + f''_{yy} \equiv 0$ in Ω. The *Poisson kernel* is the function $P : U \times \mathbb{T} \to [0, \infty)$ given by

$$P(re^{i\phi}, \theta) = \frac{1 - r^2}{2\pi(1 + r^2 - 2r\cos(\theta - \phi))} \in \Big[\frac{1 - r^2}{2\pi(1 + r)^2}, \frac{1 - r^2}{2\pi(1 - r)^2}\Big]$$

(identifying $e^{i\theta} \in \mathbb{T}$ with θ). It can be shown that if u is harmonic in $(1 + \epsilon)U$ for some $\epsilon > 0$, then

$$u(z) = \int_{\mathbb{T}} P(z, \theta) u(e^{i\theta}) d\theta =: (Pu)(z),\ z \in U.$$

In particular, with $u \equiv 1$,

$$\int_{\mathbb{T}} P(re^{i\phi}, \theta) d\phi = \int_{\mathbb{T}} P(re^{i\phi}, \theta) d\theta = 1$$

for all $r < 1$, where the first equality follows from the symmetry between ϕ and θ in the definition of P. Now fix $1 < p \le \infty$ and take $f \in L^p(\mathbb{T})$. Define

$$u(z) := (Pf)(z) = \int_{\mathbb{T}} P(z, \theta) f(\theta) d\theta.$$

It can be shown that u thus defined is harmonic. Also, we have

$$|u(z)|^p \le \Big(\int_{\mathbb{T}} |f(\theta)| P(z,\theta) d\theta \Big)^p \le \int_{\mathbb{T}} |f(\theta)|^p P(z,\theta) d\theta$$

where the last inequality follows from Proposition 2.6 applied to the measure $d\mu(\theta) = P(z,\theta)d\theta$. Hence, for $r < 1$,

$$\int_{\mathbb{T}} |u(re^{i\phi})|^p d\phi \le \int_{\mathbb{T}} \int_{\mathbb{T}} |f(\theta)|^p P(re^{i\phi},\theta) d\phi d\theta$$
$$= \int_{\mathbb{T}} |f(\theta)|^p \underbrace{\int_{\mathbb{T}} P(re^{i\phi},\theta) d\phi}_{=1} d\theta = \|f\|_p^p.$$

In summary, if $f \in L^p(\mathbb{T})$, then Pf is harmonic and $\sup_{r<1} \int_{\mathbb{T}} |(Pf)(re^{i\phi})|^p d\phi \le \|f\|_p^p$.

Conversely, if u is harmonic in U and $\sup_{r<1} \int_{\mathbb{T}} |u(re^{i\phi})|^p d\phi =: M < \infty$, then there exists an $f \in L^p(\mathbb{T})$ such that $u(z) = (Pf)(z)$, $z \in U$.

To prove the existence of such an f, let for $r < 1$, $f_r(\theta) = u(re^{i\theta})$. Then $\|f_r\|_p \le M^{1/p}$, so by Alaoglu's Theorem, there exist $r_j \uparrow 1$ such that $f_{r_j} \to f$ weakly*, i.e. $f_{r_j} \to f$ a.e., for some $f \in L^p$. The function $u(r_j z)$ is harmonic on $r_j^{-1}U$ so by the above,

$$u(r_j z) = \int_{\mathbb{T}} P(z,\theta) f_{r_j}(\theta) d\theta.$$

Since the $P(z,\theta)$ on the right hand side is bounded by $(1-r_j^2)/(2\pi(1-r_j)^2)$ and the f_{r_j}'s are bounded in L^p and $p > 1$, it follows from a version of the DCT, to be proved below, that the right hand side converges to $\int_{\mathbb{T}} P(z,\theta) f(\theta) d\theta$. The left hand side converges to $u(z)$, so we are done. $\square$

Here is the promised version of the DCT.

Theorem 7.2 *(***The Dominated Convergence Theorem***) Let μ be finite. Assume that $f_n \to f$ a.e. and that $\{f_n\}$ is uniformly integrable, i.e. for every $\epsilon > 0$, there exists $K < \infty$ such that* $\sup_n \int |f_n| \chi_{\{x : |f_n(x)| > K\}} d\mu < \epsilon$*. Then $\int f_n d\mu \to \int f d\mu$.*

Proof. Fix $\epsilon > 0$ and K as in the theorem. Then

$$\limsup_n \Big| \int f_n - \int f \Big| \le \epsilon + \int_{\{|f_n(x)| \le K\}} (f_n - f) = \epsilon$$

by the ordinary DCT. $\square$

Next we show how this version of the DCT applies to the above. Recall that $p > 1$, so that the conjugate exponent q is finite. Suppose that $\{f_n\}$ is bounded in L^p: $\sup_n \|f_n\|_p \leq M < \infty$. By Hölder,

$$\int |f_n|\chi_{\{|f_n(x)|>K\}} d\mu \leq \|f_n\|_p \mu\{|f_n(x)| > K\}^{1/q} = \|f_n\|_p \mu\{|f_n(x)|^p > K^p\}^{1/q}$$

$$\leq M\Big(\frac{M^p}{K^p}\Big)^{1/q} < \epsilon$$

for large enough K. Note that this line of reasoning does not work for $p = 1$.

8 The Riesz Representation Theorem

We recall some topological facts. Proofs of these will be given in a later appendix. A topological space X is said to be *locally compact* if for every $x \in X$, there is a compact set K such that $x \in K^o$. One says that X is *Hausdorff* if for all distinct $x, y \in X$ there are disjoint open sets V and W such that $x \in V$ and $y \in W$. When X is locally compact and Hausdorff, we write for short that X is LCH. Any LCH space X has the property that for any $x \in X$ and open set $U \ni X$, there is a compact set K such that $x \in K^o \subset K \subset U$.

Here are some functions spaces that will be extensively considered. Let Y be a topological space (usually either $\mathbb{C}$ or $\mathbb{R}$ or some subset of one of them).

- $C(X, Y) = \{f : X \to Y : f$ continuous$\}$.
- $C_0(X, Y) = \{f \in C(X, Y) : \{x : |f(x)| \geq \epsilon\}$ compact for all $\epsilon > 0\}$.
- $C_c(X, Y) = \{f \in C(X, Y) : \exists K : f|_{K^c} \equiv 0$ and K compact$\}$.

Clearly $C_c(X, Y) \subseteq C_0(X, Y) \subseteq C(X, Y)$ with equalities if X is compact. Special cases are $c_{00} \subseteq c_0 \subseteq l^1 \subseteq l^p \subseteq l^\infty \subseteq C(\mathbb{N}, \mathbb{C})$. These spaces will all be regarded as normed vector spaces, equipped with the uniform norm:

$$\|f\|_u = \sup_x |f(x)|.$$

Lemma 8.1 *(**Urysohn's Lemma**) Let X be LCH. If K is compact, U is open and $K \subset U$, then there exists a continuous function $f \in C_c(X, [0, 1])$ such that $f \equiv 1$ on K and $f \equiv 0$ on $X \setminus U$.*

Let X be LCH and let μ be a measure on $\mathcal{B}_X$. One says that μ is *outer regular* on $E \in \mathcal{B}_X$ if

$$\mu(E) = \inf\{\mu(U) : U \supseteq E, U \text{ open}\}.$$

One says that μ is *inner regular* on E if

$$\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ compact}\}.$$

If μ is outer regular on all $E \in \mathcal{B}_X$, one says that μ is outer regular and analogously for inner regularity. If μ is both outer and inner regular, then we say that μ is *regular*. A well-known example of a regular measure is Lebesgue measure.

If $\mu(K) < \infty$ for all compact K, μ is outer regular and μ is inner regular on all open sets, then μ is said to be a *Radon measure* (RM).

If μ is a RM and $f \in C_c(X, \mathbb{C})$, then $\mu\{x : |f(x)| > 0\} < \infty$. Hence $f \in L^1(\mu)$ and we get

$$C_c(X, \mathbb{C}) \subseteq L^1(\mu).$$

Hence we can define the linear functional $f \to \int f \, d\mu$, $f \in C_c(X, \mathbb{C})$. The given linear functional is *positive* in the sense that the result is positive whenever $f \geq 0$. Our first version of Riesz theorem will state that every positive linear functional, I, on $C_c(X, \mathbb{R})$ can be written this way, i.e. that there exists a RM μ such that $I(f) = \int f \, d\mu$, $f \in C_c(X, \mathbb{R})$. (Note that a positive linear functional I on $C_c(X, \mathbb{R})$ must be real-valued, since $I(f) = I(f^+) - I(f^-)$.) Since we will for the time being, work with $C_c(X, \mathbb{R})$ rather than $C_c(X, \mathbb{C})$, we write $C_c(X)$ for $C_c(X, \mathbb{R})$ until further notice. We will later come back to an extension to all linear functionals on $C_0(X, \mathbb{C})$.

Recall that the *support*, $\operatorname{supp}(f)$ of the function f is the closure of $\{x : f(x) \neq 0\}$. For an open set $U \subseteq X$, we write $f < U$ if $0 \leq f \leq 1$ and $\operatorname{supp}(f) \subset U$. Note that $f < U$ is a slightly stronger statement than $f \leq \chi_U$. We will often make use of a similarly stronger version of Urysohn's Lemma than stated above. Namely for $K \subset U$ there exists an $f \in C_c(X, [0,1])$ with $f \equiv 1$ on K and $f < U$. This follows on, for each $x \in K$, picking a compact set K_x such that $x \in K_x^o \subset K_x \subset U$. Then $\{K_x^o\}$ is an open cover of K and can hence be reduced to a finite subcover. Now apply Urysohn as above to K and the union of the sets in this subcover.

Theorem 8.2 *(***The Riesz Representation Theorem***) Let* $I : C_c(X, \mathbb{R}) \to \mathbb{R}$ *be a positive linear functional. Then there is a unique Radon measure* μ *such that*

$$I(f) = \int f \, d\mu, \; f \in C_c(X, \mathbb{R}).$$

Moreover

(*) $\mu(U) = \sup\{I(f) : f < U\}$ *for all open* U,

(**) $\mu(K) = \inf\{I(f) : f \geq \chi_K\}$ *for all compact* K.

The proof will be divided into several parts.

Proof of uniqueness and (*). Suppose $I(f) = \int f\,d\mu$, $f \in C_c(X)$, where μ is a RM. If U is open and $f < U$, then trivially $I(f) \leq \mu(U)$. If $K \subset U$ is compact, there exists by Urysohn an $f \in C_c(X)$ such that $f|_K \equiv 1$ and $f < U$. Hence $\mu(K) \leq I(f)$. Now (*) follows from the inner regularity of μ on open sets. From (*) it now follows that μ is unique on opens sets, and hence on all measurable sets by μ's outer regularity.

Proof of existence and (**). The proof will rely on several claims that will be proved afterwards.

Define the set functions

$$\mu(U) = \sup\{I(f) : f < U\},\ U \text{ open}$$

and

$$\mu^*(E) = \inf\{\mu(U) : U \supseteq E, U \text{ open}\},\ E \subseteq X.$$

Claim I. μ^* is an outer measure.

Claim II. All open sets are μ^*-measurable.

By these claims, Carathéodory's Theorem tells us that $\mu := \mu^*|_{\mathcal{B}_X}$ is a measure. This measure satisfies (*) and is outer regular by definition. (The measure μ is an extension of the set function μ above, so we allow ourselves to use the same notation for them.)

Claim III. μ satisfies (**).

From this we can quickly show that μ must be a RM. That μ is finite on compact sets follows immediately from (**). Also, if U is open and $a < \mu(U)$, then one can by (*) find $f < U$ with $I(f) > a$. Let $K = \operatorname{supp}(f) \subset U$. For any $g \geq \chi_K$ we have $g \geq f$, so $I(g) \geq I(f) > a$. Hence $\mu(K) > a$ by (**).

At this point it only remains to prove:

Claim IV. $I(f) = \int f\,d\mu$, $f \in C_c(X)$.

Before proving the claims, we need a preparatory lemma.

Lemma 8.3 *Let* K *be compact and let* $\{U_j\}_{j=1}^n$ *be an open cover of* K. *Then one can find* g_j, $j = 1, \ldots, n$, *such that* $g_j < U_j$ *and* $\sum_1^n g_j \equiv 1$ *on* K.

Proof. For each $x \in K$, $x \in U_j$ for some j. Pick a compact set N_x such that $x \in N_x^o \subset N_x \subset U_j$. Then $\{N_x^o\}_{x \in K}$ is an open cover of K. Reduce to a finite subcover $N_{x_1}, \ldots, N_{x_m}$. Let

$$F_j = \bigcup_{k: N_{x_k} \subset U_j} N_{x_k},\ j = 1, \ldots, n.$$

Then F_j is compact, so there exists $h_j \in C_c(X)$ such that $\chi_{F_j} \leq h_j < U_j$. Since $\cup_1^n F_j$ contains K, we have $\sum_1^n h_j \geq 1$ on K. Hence the open set

$$V := \{x : \sum_1^n h_j > 0\}$$

contains K. Thus we can find $f \in C_c(X)$ with $\chi_K \leq f < V$. Let $h_{n+1} = 1 - f$. Then $\sum_1^{n+1} h_j > 0$ everywhere, so that we can define

$$g_j := \frac{h_j}{\sum_1^{n+1} h_k},\ j = 1, \ldots, n.$$

Then $\mathrm{supp}(g_j) = \mathrm{supp}(h_j) \subset U_j$ and since $h_{n+1} = 0$ on K, $\sum_1^n g_j \equiv 1$ on K. $\square$

Proof of Claim I. That $\mu^*(\emptyset) = 0$ and $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$ is obvious, so it suffices to show that $\mu^*(\cup_1^\infty E_j) \leq \sum_1^\infty \mu^*(E_j)$ for any $E_j \subseteq X$. For this it suffices, by the definition of μ^*, to show that $\mu(\cup_1^\infty U_j) \leq \sum_1^\infty \mu(U_j)$ for open sets U_j. By the definition of μ, this amounts to showing that $I(f) \leq \sum_1^\infty \mu(U_j)$ for any $f \in C_c(X)$ with $f < \cup_1^\infty U_j$. Let $K = \mathrm{supp}(f)$. Since K is compact, there is an n such that $K \subset \cup_1^n U_j$. By the lemma above, there exist g_j with $g_j < U_j$ and $\sum_1^n g_j \equiv 1$ on K. However, then $f = \sum_1^n f g_j$ and $f g_j < U_j$, so

$$I(f) = \sum_1^n I(f g_j) \leq \sum_1^n \mu(U_j) \leq \sum_1^\infty \mu(U_j).$$

$\square$

Proof of Claim II. We want to show that for an open set U,

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \cap U^c).$$

Fix $\epsilon > 0$. If E is open, then $E \cap U$ is open, so we can find $f < E \cap U$ such that $\mu^*(E \cap U) = \mu(E \cap U) < I(f) + \epsilon$. In the same way there exists $g < E \setminus \mathrm{supp}(f)$ such that $\mu(E \setminus \mathrm{supp}(f)) < I(g) + \epsilon$. Since $f + g < E$,

$$\mu^*(E) = \mu(E) \geq I(f) + I(g) > \mu(E \cap U) + \mu(E \setminus \mathrm{supp}(f)) - 2\epsilon$$

$$\geq \mu^*(E \cap U) + \mu^*(E \cap U^c) - 2\epsilon.$$

This proves the desired inequality for open sets. For general E, pick open V with $\mu(V) < \mu^*(E) + \epsilon$. Then by the above,

$$\mu^*(E) \geq \mu(V) - \epsilon \geq \mu^*(V \cap U) + \mu^*(V \cap U^c) - \epsilon$$

$$\geq \mu^*(E \cap U) + \mu^*(E \cap U^c) - \epsilon.$$

□

Proof of Claim III. Fix a compact set K, $\epsilon > 0$ and $f \geq \chi_K$. Let $U_\epsilon = \{x : f(x) > 1 - \epsilon\}$. Then U_ϵ is open and contains K. Hence, for any $g < U_\epsilon$, $g \leq f/(1-\epsilon)$, so that $I(g) \leq I(f)/(1-\epsilon)$. Thus $\mu(K) \leq \mu(U_\epsilon) \leq I(f)/(1-\epsilon)$. Since ϵ is arbitrary, $\mu(K) \leq \inf\{I(f) : f \geq \chi_K\}$.

For the reverse inequality it suffices, by the outer regularity of μ to show that $\inf\{I(f) : f \geq \chi_K\} \leq \mu(U)$ for any open U that contains K. However, by Urysohn, there is an f with $\chi_K \leq f < U$ and for this f, $I(f) \leq \mu(U)$. □

Proof of Claim IV. By linearity, we may assume without loss of generality that $0 \leq f \leq 1$. Fix a large integer N. Let $K_0 = \operatorname{supp}(f)$ and $K_j = \{x : f(x) \geq j/N\}$, $j = 1, \ldots, N$. Let

$$f_j(x) = \begin{cases} 0, & x \notin K_{j-1} \\ f(x) - \frac{j-1}{N}, & x \in K_{j-1} \setminus K_j \\ \frac{1}{N}, & x \in K_j \end{cases}$$

Then $f \in C_c(X)$, $f = \sum_1^N f_j$ and $\chi_{K_j}/N \leq f_j \leq \chi_{K_{j-1}}/N$. Integrating gives

$$\frac{1}{N}\mu(K_j) \leq \int f_j \, d\mu \leq \frac{1}{N}\mu(K_{j-1}).$$

On the other hand, if $U \supset K_{j-1}$ and U is open, then $Nf_j < U$ so that $NI(f_j) \leq \mu(U)$. Since μ is outer regular, $NI(f_j) \leq \mu(K_{j-1})$. By (**), $\mu(K_j) \leq NI(f_j)$. Thus

$$\frac{1}{N}\mu(K_j) \leq I(f_j) \leq \frac{1}{N}\mu(K_{j-1}).$$

We get

$$\Big|I(f) - \int f \, d\mu\Big| \leq \sum_1^N \Big|I(f_j) - \int f_j \, d\mu\Big| \leq \sum_1^N \frac{1}{N}\Big(\mu(K_{j-1}) - \mu(K_j)\Big)$$

$$\leq \frac{1}{N}\Big(\mu(K_0) - \mu(K_N)\Big) \leq \frac{1}{N}\mu(\mathrm{supp}(f)) \to 0$$

as $N \to \infty$. □

We now have a characterization of positive linear functionals in $C_c(X, \mathbb{R})^*$. In the end we want a characterization of all linear functionals in $C_0(X, \mathbb{C})^*$. For that we first need some approximation results. From now on the notation $C_c(X)$ will be understood to mean $C_c(X, \mathbb{C})$. By definition, a RM is inner regular on all open sets. In fact, more is true:

Proposition 8.4 *A Radon measure μ is inner regular on all σ-finite sets.*

Proof. Fix $\epsilon > 0$. If $\mu(E) < \infty$, then we can find an open set $U \supseteq E$ with $\mu(U) < \mu(E) + \epsilon$. We can also pick compact $F \subset U$ with $\mu(K) > \mu(U) - \epsilon$. Since $\mu(U \setminus E) < \epsilon$ we can also find open $V \supseteq U \setminus E$ with $\mu(V) < \epsilon$. The set $K := F \setminus V$ is compact and

$$\mu(K) \geq \mu(F) - \mu(V) > \mu(E) - 2\epsilon.$$

The extension to σ-finite E is straightforward. □

Proposition 8.5 *If μ is a σ-finite Radon measure and $\epsilon > 0$, then for all $E \in \mathcal{B}_X$, there exist a closed set F and an open set U such that $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \epsilon$*

Proof. Write $E = \cup_j E_j$ where $\mu(E_j) < \infty$ for all j. Pick open $U_j \supseteq E_j$ such that $\mu(U_j) < \mu(E_j) + \epsilon 2^{-(j+1)}$. With $U = \cup_j U_j$, we get $\mu(U) < \mu(E) + \epsilon/2$, so $\mu(U \cap E^c) < \epsilon/2$. Repeat the procedure for E^c to get open $V \supseteq E^c$ with $\mu(V) < \mu(E^c) + \epsilon/2$, so $\mu(V \cap E) < \epsilon/2$. Let $F = V^c$. Then

$$\begin{aligned}\mu(U \setminus F) = \mu(U \cap V) &\leq \mu(U \cap V \cap E^c) + \mu(U \cap V \cap E)\\ &\leq \mu(U \cap E^c) + \mu(V \cap E) < \epsilon.\end{aligned}$$

□

An immediate consequence or Proposition 8.5 is that for any measurable set E one can find an F_γ-set A and G_δ-set B such that $A \subseteq E \subseteq B$ and $\mu(A) = \mu(E) = \mu(B)$.

A set $A \subseteq X$ is said to be *σ-compact* if it is the union of a countable family of compact sets. For example, if X is second countable (on top of being LCH), then every open set is σ-compact.

Theorem 8.6 *If every open set is σ-compact, then any measure μ which is finite on compact sets is regular, and hence a Radon measure.*

Proof. Since μ is finite on compact sets, $C_c(X) \subseteq L^1(\mu)$ and hence the operation $I : C_c(X) \to \mathbb{R}$ given by

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

is a well defined positive linear functional. Hence there is a RM ν such that $I(f) = \int f \, d\nu$, $f \in C_c(X)$.

Let U be open. By assumption, we can write $U = \cup_j K_j$, there the K_j's are compact and increasing. By Urysohn, there are $f_j \in C_c(X)$ such that $\chi_{K_j} \leq f_j < U$ and we may pick the f_j's so that $f_1 \leq f_2 \leq \ldots$. Then by the MCT

$$\mu(U) = \lim_n \int f_n \, d\mu = \lim_n \int f_n \, d\nu = \nu(U).$$

Now pick an arbitrary $E \in \mathcal{B}_X$. Pick, according to Proposition 8.5, open V and closed F with $F \subseteq E \subseteq V$ and $\nu(F) + \epsilon/2 > \nu(E) > /\nu(V) - \epsilon/2$. Then, since $V \setminus F$ is open,

$$\mu(V \setminus F) = \nu(V \setminus F) < \epsilon.$$

Hence $\mu(V) < \mu(F) + \epsilon \leq \mu(E) + \epsilon$, so μ is outer regular. Also, since X itself is open, we can write $X = \cup_j K_j$ for increasing compact K_j's. Thus $\mu(F) = \lim_j \mu(F \cap K_j)$, so for large enough j,

$$\mu(F \cap K_j) > \mu(F) - \epsilon > \mu(V) - 2\epsilon \geq \mu(E) - 2\epsilon.$$

Since $F \cap K_j$ is compact, this proves inner regularity. □

Proposition 8.7 *(***Approximation of measurable functions.***) If μ is a Radon measure, then $C_c(X, \mathbb{C})$ is dense in $L^p(\mu)$ for every $p \in [1, \infty)$.*

Proof. By Proposition 2.5, the set of simple functions is dense in L^p, so it suffices to show that for any measurable E of finite measure and $\epsilon > 0$, there is an $f \in C_c(X, \mathbb{C})$ with $\|\chi_E - f\|_p^p < \epsilon$. By Proposition 8.4, μ is inner regular on E, so we can find a compact K and an open U such that $K \subseteq E \subseteq U$ and $\mu(U \setminus K) < \epsilon$. By Urysohn, there is an $f \in C_c(X, \mathbb{C})$ such that $\chi_K \leq f \leq \chi_U$. Hence

$$\|\chi_E - f\|_p^p \leq \mu(U \setminus K) < \epsilon.$$

□

Theorem 8.8 *(***Luzin's Theorem***) Let μ be a Radon measure, $f : X \to \mathbb{C}$ measurable and $\epsilon > 0$. If $\mu(E) < \infty$, where $E = \{x : f(x) \neq 0\}$, then there exists a $\phi \in C_c(X)$ such that $\mu\{x : f(x) \neq \phi(x)\} < \epsilon$. If f is bounded, then ϕ can be chosen so that $\|\phi\|_u \leq \|f\|_u$.*

Proof. Recall Egoroff's Theorem: if μ is a finite measure and $f_n \to f$ a.e., then there is a set A with $\mu(A) < \epsilon$ and $f_n \to f$ uniformly on A^c.

Assume first that f is bounded. Then $f \in L^1$, so by the above proposition there exist functions $g_n \in C_c(X)$ such that $g_n \to f$ in L^1. Hence, by exercise 18, one can find $g_n \in C_c(X)$ with $g_n \to f$ a.e. By Egoroff, there is an $A \subseteq E$ such that $\mu(E \setminus A) < \epsilon/3$ and $g_n \to f$ uniformly on A^c. By Proposition 8.4 there is a compact K and an open U such that $K \subseteq A \subseteq E \subseteq U$, $\mu(A \setminus K) < \epsilon/3$ and $\mu(U \setminus E) < \epsilon/3$. Since the g_n's are continuous and $g_n \to f$ uniformly on K, f must be continuous on K. By Tietze's Extension Theorem, there exists a continuous function g with $g = f$ on K and $\mathrm{supp}(g) \subset U$. Also, $\mu\{x : \phi(x) \neq f(x)\} \leq \mu(U \setminus K) < \epsilon$.

Finally let $h : \mathbb{C} \to \mathbb{C}$ be given by $h(z) = z$ when $|z| < \|f\|_u$ and $h(z) = \|f\|_u \mathrm{sgn} z$ when $|z| \geq \|f\|_u$. Then h is continuous and $\phi = h \circ g$ satisfies also $\|\phi\|_u \leq \|f\|_u$.

The extension to unbounded f follows on picking M so that $\mu\{x : |f(x)| > M\} < \epsilon/2$ and applying the above to $f\xi_{\{x:|f(x)| \leq M\}}$. □

Now we start to build up to the full version of Riesz theorem. First we claim that $C_0(X) = \overline{C_c(X)}$: if $f \in C_0(X)$, let $K_n = \{x : |f(x)| \geq 1/n\}$ which is a compact set. By Urysohn there is a $g_n \in C_c(X)$ such that $g_n \geq \chi_{K_n}$. Then $g_n f \in C_c(X)$ and $\|g_n f - f\|_u \leq 1/n$.

Let I be a bounded positive linear functional on $C_c(X)$. For $f \in C_0(X)$ pick $f_n \in C_c(X)$ so that $\|f_n - f\|_u \to 0$. Since I is bounded, this implies that $\{I(f_n)\}$ is Cauchy and hence convergent. We can thus continuously extend I to $C_0(X)$: $I(f) = \lim_n I(f_n)$. Moreover, by Riesz, $I(f) = \int f \, d\mu$, $f \in C_c(X)$, for a RM μ. Thus, saying that I is bounded, i.e. $\|I\| < \infty$, is equivalent to $\sup\{\int g \, d\mu : g \in C_c(X), \|g\|_u \leq 1\} < \infty$, which in turn is equivalent to $\mu(X) < \infty$, by (*) applied to X. We also see that $\|I\| = \mu(X)$ and since μ is finite it follows from the DCT that

$$I(f) = \lim_n \int f_n \, d\mu = \int f \, d\mu.$$

In summary: every positive bounded linear functional $I \in C_0(X)^*$ corresponds to a finite Radon measure μ via the relation

$$I(f) = \int f\, d\mu,\ f \in C_0(X).$$

Next we want a complete characterization of $C_0(X)^*$ and not only for the positive linear functionals there.

Theorem 8.9 *(***Jordan decomposition of linear functionals***) For any real-valued $I \in C_0(X,\mathbb{R})^*$, there are positive linear functionals I^+ and I^- such that $I = I^+ - I^-$.*

Proof. Let for all nonnegative $f \in C_0(X,\mathbb{R})$

$$I^+(f) = \sup\{I(g) : 0 \leq g \leq f\}.$$

We first show that I^+ is bounded and linear for nonnegative f's. For any $0 \leq g \leq f$, $|I(g)| \leq \|I\|\|g\|_u \leq \|I\|\|f\|_u$, so $|I^+(f)| \leq \|I\|\|f\|_u$.

Clearly $I^+(cf) = cI^+(f)$ for $c \geq 0$. We also need to show that $I^+(f_1 + f_2) = I^+(f_1) + I^+(f_2)$, $f_1, f_2 \geq 0$. For this, observe that on the one hand, whenever $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$, $I^+(f_1 + f_2) \geq I(g_1 + g_2) + I(g_1) + I(g_2)$ and hence

$$I^+(f_1 + f_2) \geq I^+(f_1) + I^+(f_2).$$

On the other hand, if $0 \leq g \leq f_1 + f_2$, let $g_1 = \min(f_1, g)$ and $g_2 = g - g_1$. Then $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$. Hence

$$I^+(f_1) + I^+(f_2) \geq I^+(f_1 + f_2).$$

Next, extend I^+ to all $f \in C_0(X,\mathbb{R})$ by

$$I^+(f) = I^+(f^+) - I^+(f^-).$$

To see that this is well defined, suppose that f also can be written as $g - h$, $g, h \geq 0$. Then $f^+ + h = g + f^-$ so by the above linearity, $I^+(f^+) + I^+(h) = I^+(g) + I^+(f^-)$, i.e. $I^+(f^+) - I^+(f^-) = I^+(g) - I^+(h)$. The functional I^+ is obviously linear and

$$|I^+(f)| \leq \max(|I^+(f^+)|, |I^+(f^-)|) \leq \|I\| \max(\|f^+\|_u, \|f^-\|_u) = \|I\|\|f\|_u$$

so that $\|I^+\| \leq \|I\|$. Finally let $I^- = I^+ - I$, which is positive by the definition of I^+. □

Assume now that $I \in C_0(X)^*$ and write J for the restriction of I to $C_0(X, \mathbb{R})$. Then, writing $f = g + ih$ for $f \in C_0(X)$, we have

$$I(f) = J(g) + iJ(h).$$

Let also $J_1 = \Re J$ and $J_2 = \Im J$. By the decomposition lemma, we can write $J_k = J_k^+ - J_k^-$ for positive functionals J_k^+ and J_k^-, $k = 1, 2$. By Riesz there are unique RM's μ_1, μ_2 ,μ_3 and μ_4 such that $J_1^+(f) = \int f \, d\mu_1$, $J_1^-(f) = \int f \, d\mu_2$, $J_2^+ = \int f \, d\mu_3$ and $J_2^+(f) = \int f \, d\mu_4$, $f \in C_0(X, \mathbb{R})$. Putting this together, letting μ denote the complex measure $\mu_1 - \mu_2 + i(\mu_3 - \mu_4)$,

$$I(f) = \int f \, d\mu, \ f \in C_0(X).$$

This already gives an extended version of Riesz' theorem, but one can in fact do more. A complex measure μ is said to be Radon if all its parts are Radon measures. Let

$$M(X) := \{\mu : \mathcal{B}_X \to C : \mu \text{ finite complex Radon measure}\}.$$

Make $M(X)$ into a NVS using the norm $\|\mu\| = |\mu|(X)$. Recall that the measure $|\mu|$, the total variation of μ is defined as $|\mu| = \mu_1 + \mu_2 + \mu_3 + \mu_4$. Equivalently, for any measure ν such that $\mu_k \ll \nu$, $k = 1, 2, 3, 4$,

$$|\mu|(E) = \int_E |f| \, d\nu, \ E \in \mathcal{B}_X$$

where f is a Radon-Nikodym derivative $d\mu/d\nu$. From this is follows that for two complex measures μ and μ', $|\mu + \mu'| \le |\mu| + |\mu'|$. This shows that $\| \cdot \|$ is indeed a norm.

Clearly $|\mu|$ is Radon if all its parts are Radon. Conversely, if $|\mu|$ is Radon and finite, then, for $E \in \mathcal{B}_X$, there is an open U and a compact K such that $K \subseteq E \subseteq U$ and $|\mu|(U \setminus K) < \epsilon$. This follows from combining Propositions 8.4 and 8.5. Hence $\mu_k(U \setminus K) < \epsilon$ for all parts of μ. This proves that μ_k is regular and in particular Radon. In summary: a finite complex measure μ is Radon iff $|\mu|$ is Radon.

Theorem 8.10 *(***Riesz Representation Theorem, complex version***) For $\mu \in M(X)$, let $I_\mu(f) = \int f \, d\mu$, $f \in C_0(X)$. Then the map $\mu \to I_\mu$ is an isometric isomorphism from $M(X)$ to $C_0(X)^*$.*

Proof. That the map is injective is obvious and surjectivity was shown above. Linearity is also obvious, so it remains to prove that $\|\mu\| = \|I_\mu\|$. However

$$|I_\mu(f)| = |\int f\,d\mu| \leq \int |f| d|\mu| \leq |\mu|(X)\|f\|_u$$

so $\|I_\mu\| \leq \|\mu\|$. For the reverse inequality, let $h = d\mu/d|\mu|$ and fix $\epsilon > 0$. By Luzin's Theorem there is an $f \in C_c(X)$ with $\|f\|_u \leq 1$ and $f = \overline{h}$ on E^c where $|\mu|(E) < \epsilon/2$. Since $|h| = 1$, we get

$$\|\mu\| = \int |h|^2 d|\mu| = \int \overline{h} h\,d|\mu| = \int \overline{h}\,d\mu$$

$$\leq \left|\int (\overline{h} - f) d\mu\right| + \left|\int f\,d\mu\right| \leq 2|\mu|(E) + \left|\int f\,d\mu\right| < \epsilon + |I_u(f)|.$$

Hence $\|\mu\| \leq \|I_\mu\|$. □

Corollary 8.11 *If X is compact, then $M(X)$ and $C(X)^*$ are isometrically isomorphic.*

9 Vague convergence of measures

In this section we study weak* convergence on $M(X) \cong C_0(X)^*$ in a bit more detail. Since weak* convergence of measures is usually referred to as weak convergence in probability theory, which is a stronger concept here (since $C_0(X)$ is usually not reflexive), a common compromise is to refer to weak* convergence as *vague* convergence. Hence, to say that $\mu_n \to \mu$ vaguely means the same as saying that $\mu_n \to \mu$ weakly*, i.e.

$$I_{\mu_n}(f) \to I_\mu(f),\ f \in C_0(X)$$

i.e.

$$\int f\,d\mu_n \to \int f\,d\mu,\ f \in C_0(X).$$

In probability theory, if μ_n is the distribution of the random variable ξ_n and μ is the distribution of μ, i.e. $\mu(E) = \mathbb{P}(\xi \in E)$, $E \in \mathcal{B}_\mathbb{R}$ and analogously for μ_n, and $\mu_n \to \mu$ vaguely, then this is also referred to as $\xi_n \to \xi$ in distribution.

For a positive measure in $M(\mathbb{R})$, write $F(x) = F_\mu(X) = \mu(-\infty, x]$, $x \in \mathbb{R}$. The following result is important, in particular in probability theory.

Proposition 9.1 *Assume that $\mu_n, \mu \in M(\mathbb{R})$ are positive measures such that $\sup_n \|\mu_n\| < \infty$. Then $\mu_n \to \mu$ vaguely iff $F_n(x) \to F(x)$ for every x at which F is continuous.*

Proof. We prove only the backwards implication. We want to show that $\int_{\mathbb{R}} f\, d\mu_n \to \int_{\mathbb{R}} f\, d\mu$ for all $f \in C_0(\mathbb{R})$. Assume first that f has compact support and is continuously differentiable. Since F is increasing, F is continuous at all but at most countably many points. In particular $F_n \to F$ a.e. By assumption, the F_n's are uniformly bounded. Hence integration by parts, the DCT and the fact that f has compact support imply that

$$\int_{\mathbb{R}} f\, d\mu_n = \int_{\mathbb{R}} f\, dF_n = \int_{\mathbb{R}} f'(x) F_n(x) dx$$

$$\to \int_{\mathbb{R}} f'(x) F(x) dx = \int_{\mathbb{R}} f\, d\mu.$$

By the Stone-Weierstrass Theorem, the family, S, of continuously differentiable functions with compact support, is dense in $C_0(X)$. Fix $\epsilon > 0$ and let $M = \max(\|\mu\|, \sup_n \|\mu_n\|)$. For $f \in C_0(X)$ pick $g \in S$ such that $\|f - g\|_u < \epsilon/(3M)$. By the above $|\int_{\mathbb{R}} g\, d\mu_n - \int_{\mathbb{R}} g\, d\mu| < \epsilon/3$ for sufficiently large n. An application of the triangle inequality then gives

$$\Big| \int_{\mathbb{R}} f\, d\mu_n - \int_{\mathbb{R}} f\, d\mu \Big| < \epsilon.$$

□

10 Some useful inequalities

Theorem 10.1 *(**Markov's/Chebyshev's inequality**) Let $f \in L^p(\mu)$ and let $E_\alpha = \{x : |f(x)| > \alpha\}$. Then*

$$\mu(E_\alpha) \leq \frac{\|f\|_p^p}{\alpha^p}.$$

Proof.

$$\|f\|_p^p = \int_X |f|^p d\mu \geq \int_{E_\alpha} |f|^p d\mu \geq \alpha^p \mu(E_\alpha).$$

□

Theorem 10.2 *(***Boundedness of integral operators***) Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be two σ-finite measure spaces and let $f : X \times Y \to \mathbb{C}$ be $\mathcal{M}\otimes\mathcal{N}$-measurable. Let $p \in [1,\infty]$ and let $T : L^p(\nu) \to L^p(\mu)$ be the linear operator given by*

$$(Tf)(x) = \int_Y K(x,y)f(y)\nu(dy),\ f \in L^p(\nu).$$

If there exists a constant $C < \infty$ such that

$$\int_X |K(x,y)|\mu(dx) \leq C$$

for a.e. $y \in Y$ and

$$\int_Y |K(x,y)|\nu(dy) \leq C$$

for a.e. $x \in X$, then the integral in the definition of T converges absolutely and

$$\|Tf\|_{L^p(\mu)} \leq C\|f\|_{L^p(\nu)}.$$

In particular, T is a well defined bounded linear operator.

Proof. We do the case $1 < p < \infty$, leaving the easier cases $p = 1$ and $p = \infty$ as an exercise. Let q be the conjugate exponent of p. We have

$$\begin{aligned}\int |K(x,y)f(y)|\nu(dy) &= \int |K(x,y)|^{1/q}\Big(|K(x,y)^{1/p}||f(y)|\Big)\nu(dy)\\ &\leq \Big(\int |K(x,y)|\nu(dy)\Big)^{1/q}\Big(\int |K(x,y)||f(y)|^p\nu(dy)\Big)^{1/p}\\ &\leq C^{1/q}\Big(\int |K(x,y)||f(y)|^p\nu(dy)\Big)^{1/p}\end{aligned}$$

where the first inequality is Hölder and the second inequality follows from assumption. By Tonelli's Theorem

$$\begin{aligned}\int_X \Big(\int_Y |K(x,y)f(y)|\nu(dy)\Big)^p\mu(dx) &\leq C^{p/q}\int_Y |f(y)|^p\int_X |K(x,y)|\mu(dx)\nu(dy)\\ &\leq C^{p/q+1}\int_Y |f(y)|^p\nu(dy) < \infty\end{aligned}$$

where the first inequality follows from the above, the second from assumption and the third from the fact that $f \in L^p(\nu)$. Hence $\int_Y |K(x,y)f(y)|\nu(dy)$ converges for a.e. x as desired. Also,

$$\int_X |(Tf)(x)|^p \mu(dx) \leq \int_X \Big(\int_Y |K(x,y)f(y)|\nu(dy)\Big)^p \mu(dx) \leq C^{p/q+1}\|f\|_p^p$$

so $\|Tf\|_p \leq C\|f\|_p$. □

Theorem 10.3 *(***Minkowski's inequality for integrals***) Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite, $1 \leq p < \infty$ and f a $\mathcal{M} \otimes \mathcal{N}$-measurable nonnegative function. Then*

$$\Big[\int_X \Big(\int_Y (f(x,y)\nu(dy)\Big)^p \mu(dx)\Big]^{1/p} \leq \int_Y \Big(\int_X f(x,y)^p \mu(dx)\Big)^{1/p} \nu(dy).$$

In other words

$$\| \int_Y f(\cdot,y)\nu(dy)\|_{L^p(\mu)} \leq \int_Y \|f(\cdot,y)\|_{L^p(\mu)}\nu(dy).$$

Proof. For $p = 1$ this is an equality that follows immediately from Tonelli's Theorem, so assume $p > 1$ and let q be the conjugate exponent. If $g \in L^q(\mu)$ and $\|g\|_q = 1$, then

$$\int_X \Big(\int_Y f(x,y)\nu(dy)\Big)|g(x)|\mu(dx) = \int_Y \int_X f(x,y)|g(x)|\mu(dx)\nu(dy)$$

$$\leq \int_Y \Big(\int_X f(x,y)^p \mu(dx)\Big)^{1/p} \nu(dy)$$

where the inequality follows from Hölder applied to the inner integral, keeping in mind that $\|g\|_q = 1$. Now use Theorem 3.3, the reverse of Hölder, with g as f, q as p and $\int_Y f(\cdot,y)\nu(dy)$ as g. □

11 Appendix on topology

The aim of this appendix is to tie up the "loose ends" from the section on Riesz Representation Theorem, by proving the topological facts we used there.

Let X and Y be topological spaces and assume that the topology on Y is generated by $\mathcal{E} \subset \mathcal{P}(Y)$. Let $f : X \to Y$. Then f is continuous iff $f^{-1}(V)$ is open in X for every $V \in \mathcal{E}$. The forward implication is trivial. The backward implication follows from that inverse images commute with the elementary set operations.

A topological space is said to be *normal* if for any disjoint closed sets A and B, there are disjoint open sets U and V with $U \supset A$ and $V \supset B$.

Lemma 11.1 *Let X be normal and A and B closed and disjoint subsets. Let $I = \{k2^{-n} : n \in \mathbf{N}, k = 1, 2, \ldots, 2^n\}$. Then there exist open sets U_r, $r \in I$ such that $A \subseteq U_r \subseteq B^c$ for all r and $\overline{U_r} \subset U_s$ whenever $r < s$.*

Proof. Let $U_1 = B^c$. Pick open disjoint $V \supset A$ and $W \supset B$ and let $U_{1/2} = V$. Then

$$A \subseteq U_{1/2} \subset \overline{U^{1/2}} \subseteq W^c \subseteq B^c = U_1.$$

Now use induction on n: suppose that we have sets U_r with the desired properties for $r \in \{k2^{-n} : n = 1, \ldots, N-1, \, k = 1, \ldots, 2^n\}$ and let $r = i2^{-N}$ for an odd i, i.e. $r = (2j+1)2^{-N}$ for some $j \in \{0, \ldots, 2^{N-1} - 1\}$. Then by the induction hypothesis,

$$\overline{U_{j2^{-(N-1)}}} \subset U_{(j+1)2^{-(N-1)}}.$$

By normality, there exist disjoint open sets $V \supset \overline{U_{j2^{-(N-1)}}}$ and $W \supset U^c_{(j+1)2^{-(N-1)}}$. Then

$$\overline{V} \subseteq W^c \subset U_{(j+1)2^{-(N-1)}}$$

so we can take $U_{(2j+1)2^{-N}} = V$. $\square$

Theorem 11.2 *(***Urysohn's Lemma for Normal Spaces***) Let X be normal and let A and B be disjoint closed subsets. Then there exists an $f \in C(X, [0,1])$ such that $f \equiv 0$ on A and $f \equiv 1$ on B.*

Proof. Let U_r be as in the lemma and let $U_r = X$ for $r > 1$. Let

$$f(x) = \inf\{r \in I : x \in U_r\}.$$

Trivially $f = 0$ on A, $f = 1$ on B and $f(x) \in [0,1]$ for all x. It remains to show that f is continuous. However

$$f^{-1}[0, \alpha) = \{x : f(x) < \alpha\} = \{x : x \in U_r \text{ for some } r < \alpha\} = \bigcup_{r < \alpha} U_r,$$

an open set. Similarly

$$f^{-1}(\alpha,1] = \{x : f(x) > \alpha\} = \{x \notin \overline{U_r} \text{ for some } r > \alpha\} = \bigcup_{r>\alpha} \overline{U_r}^c$$

is also open. Since the sets $[0,\alpha)$ and $(\alpha,1]$, $\alpha \in [0,1]$ generate the topology on $[0,1]$, we are done. □

If A is a closed subset of a compact space X and $\{U_\alpha\}$ is an open cover of A, then $\{U_\alpha\} \cup \{X \setminus A\}$ is an open cover of X. It follows that any closed subset of a compact space is compact.

Suppose that X is Hausdorff, K a compact subset and $x \notin K$. Then there are disjoint open sets $U \ni x$ and $V \supset K$. To see this, pick for each $y \in K$, disjoint open sets $U_y \ni x$ and $V_y \ni y$. Then $\{V_y\}_{y\in K}$ is an open cover of K and can hence be reduced to a finite subcover $V_{y_1}, \ldots, V_{y_n}$. Now take $U = \cap_1^n U_{y_k}$ and $V = \cup_1^n V_{y_k}$.

Proposition 11.3 *Every compact Hausdorff space X is normal.*

Proof. Let A and B be closed and disjoint. Since B is compact, we can for each $x \in A$ find disjoint open sets $U_x \ni x$ and $V_x \supset B$. Since A is compact, the open cover $\{U_x\}_{x\in A}$ can be reduced to a finite subcover $U_{x_1}, \ldots, U_{x_n}$. Let $U = \cup_1^n U_{x_k}$ and $V = \cap_1^n V_{x_k}$. Then U and V are open and disjoint, $U \supset A$ and $V \supset B$. □

Proposition 11.4 *Let X be a locally compact Hausdorff space. Let U be open and $x \in U$. Then there exists a compact set K such that $x \in K^o \subset K \subset U$.*

Proof. We may assume that $\overline{U}$ is compact, since if this is not the case, we can by definition of local compactness find compact F with $x \in F^o$. Then we can replace U with $U \cap F^o$ whose closure is compact.

Since $\overline{U}$ is compact, ∂U is also compact. Hence there are sets $U \ni x$ and $W \supset \partial U$ that are open in $\overline{U}$. Since $V \subseteq U$, V is open in X. We also have

$$\overline{V} \subseteq \overline{U} \setminus W \subset \overline{U}.$$

Hence $\overline{V}$ is compact, so we can take $K = \overline{V}$. □

Now suppose X is LCH, U open, K compact and $K \subset U$. For each $x \in K$, pick the compact set K_x so that $x \in K_x^o \subset K_x \subset U$. Then the K_x's constitute an open cover of K and can hence be reduced to a finite subcover $K_{x_1}, \ldots, K_{x_n}$. Let $N = \cup_1^n K_{x_i}$. Then N is compact and $K \subset N^o \subset N \subset U$. The subspace N is compact and hence normal, so by Urysohn for normal spaces there is an $f \in C(N, [0,1])$ such that $f \equiv 1$ on K and $f \equiv 0$ on ∂N. Extend to the whole of X by letting $f \equiv 0$ on $X \setminus N$. Then $f \in C(X, [0,1])$ and $f < U$. This proves Urysohn's Lemma in the form most used.

Theorem 11.5 *(***The Tietze Extension Theorem***) Let X be LCH, $K \subset X$ compact and $f \in C(K, [0,1])$. Then there exists an $F \in C(X, [0,1])$ such that $F = f$ on K and $F = 0$ outside a compact set.*

Proof. Assume first that X is normal and A and B are closed an disjoint. We claim that there exist continuous functions g_j with $0 \leq g_j \leq 2^{j-1}/3^j$ and $f - \sum_1^k g_j \leq (2/3)^k$. To see this for $j = 1$, let $C = f^{-1}[0, 1/3]$ and $D = f^{-1}[2/3, 1]$. Then C and D are closed subsets of A and since A is closed, C and D are closed in X. Hence, by Urysohn, there is a $g_1 \in C(X, [0, 1/3])$ with $g_1 = 1/3$ on D and $g_1 = 0$ on $B \cup C$.

Now suppose that we have found $g_1, \ldots, g_{n-1}$ of the desired form. Then, in the same way, there is a $g_n \in C(X, [0, 2^{n-1}/3^n])$ such that $g_n = 2^{n-1}/3^n$ on $(f - \sum_1^{n-1} g_j)^{-1}[0, 2^{n-1}/3^n]$ and $g_n = 0$ on $(f - \sum_1^{n-1} g_j)^{-1}[(2/3)^n, (2/3)^{n-1}] \cup B$. Since $(2/3)^{n-1} - 2^{n-1}/3^n = (2/3)^n$, g_n is also of the desired form.

Let $F = \sum_1^\infty g_j$. Clearly $F = f$ on A and $F = 0$ on B. Since

$$\|F - \sum_1^n g_j\|_u \leq \sum_{n+1}^\infty \left(\frac{2}{3}\right)^j \to 0,$$

the series converges uniformly and hence F is continuous.

For the general statement, apply the above with $A = K$ and $B = \partial N$, where N is compact and $K \subset N^o$, on the compact and hence normal subspace N of X. Extend by letting $F \equiv 0$ on N^c. □