Linear Analysis I: Banach and Normed Spaces

Updated: 4th November 2009

Recall what a *vector space* is. We will work with either the field of real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. To avoid repetition, we use $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$. A *norm* on a vector space V is a map $\|\cdot\| : V \to [0,\infty)$ such that $\|u\| = 0$ only when $u = 0$; $\|\lambda u\| = |\lambda|\|u\|$ for $\lambda \in \mathbb{K}$ and $u \in V$; $\|u+v\| \leq \|u\| + \|v\|$ for $u, v \in V$.

A norm induces a *metric* on V by setting $d(u,v) = \|u - v\|$. When V is *complete* for this metric, we say that V is a *Banach space*.

Theorem: Every finite-dimensional normed vector space is a Banach space. $\square$

Hölder's Inequality: For $1 < p < \infty$, let $q \in (1,\infty)$ be such that $1/p + 1/q = 1$. For $n \geq 1$ and $u, v \in \mathbb{K}^n$, we have that

$$\sum_{j=1}^n |u_j v_j| \leq \Big(\sum_{j=1}^n |u_j|^p \Big)^{1/p} \Big(\sum_{j=1}^n |v_j|^q \Big)^{1/q}.$$

Minkowski's Inequality: For $1 < p < \infty$, and $n \geq 1$, let $u, v \in \mathbb{K}^n$. Then

$$\Big(\sum_{j=1}^n |u_j + v_j|^p \Big)^{1/p} \leq \Big(\sum_{j=1}^n |u_j|^p \Big)^{1/p} + \Big(\sum_{j=1}^n |v_j|^p \Big)^{1/p}.$$

Minkowski's inequality shows that for $1 \leq p < \infty$ (the case $p = 1$ is easy) we can define a norm $\|\cdot\|_p$ on $\mathbb{K}^n$ by

$$\|u\|_p = \Big(\sum_{j=1}^n |u_j|^p \Big)^{1/p} \qquad (u = (u_1, \cdots, u_n) \in \mathbb{K}^n).$$

We can define an infinite analogue of this. Let $1 \leq p < \infty$, let ℓ^p be the space of all scalar sequences (x_n) with $\sum_n |x_n|^p < \infty$. A careful use of Minkowski's inequality shows that ℓ^p is a vector space. Then ℓ^p becomes a normed space for the $\|\cdot\|_p$ norm.

Recall that a Cauchy sequence in a normed space is bounded: if (x_n) is Cauchy then we can find N with $\|x_n - x_m\| < 1$ for all $n, m \geq N$. Then $\|x_n\| \leq \|x_n - x_N\| + \|x_N\| < \|x_N\| + 1$ for $n \geq N$, so in particular, $\|x_n\| \leq \max(\|x_1\|, \|x_2\|, \cdots, \|x_{N-1}\|, \|x_N\| + 1)$.

Theorem: For $1 \leq p < \infty$, the space ℓ^p is a Banach space.

Proof: Most completeness proofs are similar to this, so we shall prove this result in detail. Let (x_n) be a Cauchy-sequence in ℓ^p; we wish to show this converges to some vector in ℓ^p.

For each n, $x_n \in \ell^p$ so is a sequence of scalars, say $(x_k^{(n)})_{k=1}^\infty$. As (x_n) is Cauchy, for each $\epsilon > 0$ there exists N_ϵ so that $\|x_n - x_m\|_p \leq \epsilon$ for $n, m \geq N_\epsilon$.

For k fixed,

$$|x_k^{(n)} - x_k^{(m)}| \leq \Big(\sum_j |x_j^{(n)} - x_j^{(m)}|^p \Big)^{1/p} = \|x_n - x_m\|_p \leq \epsilon,$$

when $n, m \geq N_\epsilon$. Thus the scalar sequence $(x_k^{(n)})_{n=1}^\infty$ is Cauchy in $\mathbb{K}$ and hence converges, to y_k say.

Let $y = (y_k)$, so that y is a candidate for the limit of (x_n). Firstly, we check that $y \in \ell^p$. We calculate,

$$\begin{aligned}
\|y\|_p &= \lim_{K\to\infty}\Big(\sum_{k=1}^K |y_k|^p\Big)^{1/p} = \lim_{K\to\infty}\Big(\sum_{k=1}^K \lim_{n\to\infty}|x_k^{(n)}|^p\Big)^{1/p}\\
&= \lim_{K\to\infty}\lim_{n\to\infty}\Big(\sum_{k=1}^K |x_k^{(n)}|^p\Big)^{1/p} \leq \lim_{K\to\infty}\lim_{n\to\infty}\Big(\sum_{k=1}^\infty |x_k^{(n)}|^p\Big)^{1/p}\\
&= \lim_{n\to\infty}\|x_n\|_p < \infty,
\end{aligned}$$

as (x_n) is Cauchy, and hence bounded.

Finally, we check that $x_n \to y$ in ℓ^p. For $\epsilon > 0$, let $n \geq N_\epsilon$, so that

$$\begin{aligned}
\|x_n - y\|_p &= \lim_{K\to\infty}\Big(\sum_{k=1}^K |x_k^{(n)} - y_k|^p\Big)^{1/p} = \lim_{K\to\infty}\Big(\sum_{k=1}^K \lim_{m\to\infty}|x_k^{(n)} - x_k^{(m)}|^p\Big)^{1/p}\\
&= \lim_{K\to\infty}\lim_{m\to\infty}\Big(\sum_{k=1}^K |x_k^{(n)} - x_k^{(m)}|^p\Big)^{1/p} \leq \lim_{K\to\infty}\lim_{m\to\infty}\Big(\sum_{k=1}^\infty |x_k^{(n)} - x_k^{(m)}|^p\Big)^{1/p}\\
&= \lim_{m\to\infty}\|x_n - x_m\|_p \leq \epsilon,
\end{aligned}$$

as $n \geq N_\epsilon$. Hence $\|x_n - y\|_p \to 0$. □

For $p = \infty$, there are two analogies to the ℓ^p spaces. The first is arguably more natural, but we write c_0 for it. c_0 is the space of all scalar sequences (x_n) which converge to 0. We equip c_0 with the sup norm,

$$\|(x_n)\|_\infty = \sup_{n\in\mathbb{N}}|x_n| \qquad ((x_n) \in c_0).$$

This is defined, as if $x_n \to 0$, then (x_n) is bounded. Similarly, we define ℓ^∞ to be the vector space of all bounded scalar sequences, with the $\|\cdot\|_\infty$ norm. Hence c_0 is a subspace of ℓ^∞, and we can check that c_0 is closed.

Theorem: The spaces c_0 and ℓ^∞ are Banach spaces.

Proof: This will be a variant of the previous proof: it's shorter, but the "trick" is maybe harder to remember. We do the ℓ^∞ case. Again, let (x_n) be a Cauchy sequence in ℓ^∞, and for each n, let $x_n = (x_k^{(n)})_{k=1}^\infty$. For $\epsilon > 0$ we can find N such that $\|x_n - x_m\|_\infty < \epsilon$ for $n, m \geq N$. Thus, for any k, we see that $|x_k^{(n)} - x_k^{(m)}| < \epsilon$ when $n, m \geq N$. So $(x_k^{(n)})_{n=1}^\infty$ is Cauchy, and hence converges, say to $x_k \in \mathbb{K}$. Let $x = (x_k)$.

Let $m \geq N$, so that for any k, we have that

$$|x_k - x_k^{(m)}| = \lim_{n\to\infty}|x_k^{(n)} - x_k^{(m)}| \leq \epsilon.$$

As k was arbitrary, we see that $\sup_k |x_k - x_k^{(m)}| \leq \epsilon$. So, firstly, this shows that $(x - x_m) \in \ell^\infty$, and so also $x = (x - x_m) + x_m \in \ell^\infty$. Secondly, we have shown that $\|x - x_m\|_\infty \leq \epsilon$ when $m \geq N$, so $x_m \to x$ in norm. □

Bounded linear operators

Recall what a *linear* map is. A linear map $T : E \to F$ between normed spaces is *bounded* if there exists $M > 0$ such that $\|T(x)\| \leq M\|x\|$ for $x \in E$. A bounded linear map is often called an *operator*. We write $\mathcal{B}(E, F)$ for the set of operators from E to F. For the natural operations, $\mathcal{B}(E, F)$ is a vector space. We norm $\mathcal{B}(E, F)$ by setting

$$\|T\| = \sup\Big\{\frac{\|T(x)\|}{\|x\|} : x \in E, x \neq 0\Big\}.$$

This *is* a norm, and we equivalently have

$$\|T\| = \sup\{\|T(x)\| : x \in E, \|x\| \leq 1\} = \sup\{\|T(x)\| : x \in E, \|x\| = 1\}.$$

Proposition: For a linear map $T : E \to F$ between normed spaces, the following are equivalent:

1. T is continuous (for the metrics induced by the norms on E and F);
2. T is continuous at 0;
3. T is bounded. □

Theorem: Let E be a normed space, and let F be a Banach space. Then $\mathcal{B}(E, F)$ is a Banach space.

Proof: Let (T_n) be a Cauchy sequence in $\mathcal{B}(E, F)$. For $x \in E$, check that $(T_n(x))$ is Cauchy in F, and hence converges to, say, $T(x)$, as F is complete. Then check that $T : E \to F$ is linear, bounded, and that $\|T_n - T\| \to \infty$. □

We write $\mathcal{B}(E)$ for $\mathcal{B}(E, E)$. For normed spaces E, F and G, and for $T \in \mathcal{B}(E, F)$ and $S \in \mathcal{B}(F, G)$, we have that $ST = S \circ T \in \mathcal{B}(E, G)$ with $\|ST\| \leq \|S\|\|T\|$.

For $T \in \mathcal{B}(E, F)$, if there exists $S \in \mathcal{B}(F, E)$ with $ST = I_E$, the identity of E, and $TS = I_F$, then T is said to be *invertible*, and write $T = S^{-1}$. In this case, we say that E and F are *isomorphic* spaces, and that T is an *isomorphism*.

If $\|T(x)\| = \|x\|$ for each $x \in E$, we say that T is an *isometry*. If additionally T is an isomorphism, then T is an *isometric isomorphism*, and we say that E and F are *isometrically isomorphic*.

Dual Spaces

Let E be a normed vector space, and let E^* (also written E') be $\mathcal{B}(E, \mathbb{K})$, the space of bounded linear maps from E to $\mathbb{K}$, which we call *functionals*, or more correctly, *bounded linear functionals*. Notice that as $\mathbb{K}$ is complete, the above theorem shows that E^* is always a Banach space.

Theorem: Let $1 < p < \infty$, and again let q be such that $1/p + 1/q = 1$. Then the map $\ell^q \to (\ell^p)^*; u \mapsto \phi_u$, is an isometric isomorphism, where ϕ_u is defined, for $u = (u_j) \in \ell^q$, by

$$\phi_u(x) = \sum_{j=1}^{\infty} u_j x_j \qquad \big(x = (x_j) \in \ell^p\big).$$

Proof: By Holder's inequality, we see that

$$|\phi_u(x)| \leq \sum_{j=1}^{\infty} |u_j||x_j| \leq \Big(\sum_{j=1}^{\infty} |u_j|^q\Big)^{1/q} \Big(\sum_{j=1}^{\infty} |x_j|^p\Big)^{1/p} = \|u\|_q \|x\|_p.$$

So the sum converges, and hence ϕ_u is defined. Clearly ϕ_u is linear, and the above estimate also shows that $\|\phi_u\| \leq \|u\|_q$. The map $u \mapsto \phi_u$ is also clearly linear, and we've just shown that it is norm-decreasing.

Now let $\phi \in (\ell^p)^*$. For each n, let $e_n = (0, \cdots, 0, 1, 0, \cdots)$ with the 1 in the nth position. Then, for $x = (x_n) \in \ell^p$,

$$\Big\|x - \sum_{k=1}^{n} x_k e_k\Big\|_p = \Big(\sum_{k=n+1}^{\infty} |x_k|^p\Big)^{1/p} \to 0,$$

as $n \to \infty$. As ϕ is continuous, we see that

$$\phi(x) = \lim_{n\to\infty} \sum_{k=1}^{n} \phi(x_k e_k) = \sum_{k=1}^{\infty} x_k \phi(e_k).$$

Let $u_k = \phi(e_k)$ for each k. If $u = (u_k) \in \ell^q$ then we would have that $\phi = \phi_u$.

Let $N \in \mathbb{N}$, and define

$$x_k = \begin{cases} 0 & : u_k = 0 \text{ or } k > N, \\ \overline{u_k}|u_k|^{q-2} & : u_k \neq 0. \end{cases}$$

Then we see that

$$\sum_{k=1}^{\infty} |x_k|^p = \sum_{k=1}^{N} |u_k|^{p(q-1)} = \sum_{k=1}^{N} |u_k|^q,$$

as $p(q-1) = q$. So $x = (x_k) \in \ell^p$. Then, by the previous paragraph,

$$\phi(x) = \sum_{k=1}^{\infty} x_k u_k = \sum_{k=1}^{N} |u_k|^q.$$

Hence

$$\|\phi\| \geq \frac{|\phi(x)|}{\|x\|_p} = \Big(\sum_{k=1}^{N} |u_k|^q \Big)^{1-1/p} = \Big(\sum_{k=1}^{N} |u_k|^q \Big)^{1/q}.$$

By letting $N \to \infty$, it follows that $u \in \ell^q$ with $\|u\|_q \leq \|\phi\|$. So $\phi = \phi_u$ and $\|\phi\| = \|\phi_u\| \leq \|u\|_q$. Hence every element of $(\ell^p)^*$ arises as ϕ_u for some u, and also $\|\phi_u\| = \|u\|_q$. $\square$

Loosely speaking, we say that $\ell^q = (\ell^p)^*$, although we should always be careful to keep in mind the exact map which gives this.

Similarly, we can show that $c_0^* = \ell^1$ and that $(\ell^1)^* = \ell^\infty$ (the implementing isometric isomorphism is giving by the same summation formula).

Diversion: Zorn's Lemma. A *poset* is a set X with a relation $\preceq$ such that $a \preceq a$ for all $a \in X$, if $a \preceq b$ and $b \preceq a$ then $a = b$, and if $a \preceq b$ and $b \preceq c$, then $a \preceq c$. We say that $(X, \preceq)$ is *total* if for every $a, b \in X$, either $a \preceq b$ or $b \preceq a$. For a subset $S \subseteq X$, an element $a \in X$ is an *upper bound* for S if $s \preceq a$ for every $s \in S$. An element $a \in X$ is *maximal* if whenever $b \in X$ is such that $a \preceq b$, then also $b \preceq a$.

Then **Zorn's Lemma** tells us that if X is a non-empty poset such that every total subset has an upper bound, then X has a maximal element. Really this is an *axiom* which we have to assume, in addition to the usual axioms of set-theory. Zorn's Lemma is equivalent to the *axiom of choice*.

Hahn-Banach Theorem: Let E be a normed vector space, and let $F \subseteq E$ be a subspace. Let $\phi \in F^*$. Then there exists $\psi \in E^*$ with $\|\psi\| \leq \|\phi\|$ and $\psi(x) = \phi(x)$ for each $x \in F$.

Proof: We do the real case. An "extension" of ϕ is a bounded linear map $\phi_G : G \to \mathbb{R}$ such that $F \subseteq G \subseteq E$, $\phi_G(x) = \phi(x)$ for $x \in F$, and $\|\phi_G\| \leq \|\phi\|$. A Zorn's Lemma argument shows that a maximal extension $\phi_G : G \to \mathbb{R}$ exists. We shall show that if $G \neq E$, then we can extend ϕ_G, a contradiction.

Let $x_0 \notin G$, so an extension ϕ_0 of ϕ_G to the linear span of G and x_0 must have the form

$$\phi_0(x + a x_0) = \phi_G(x) + a\alpha_0 \qquad (x \in G, a \in \mathbb{R}),$$

for some $\alpha_0 \in \mathbb{R}$. Under this, ϕ_0 is linear and extends ϕ_G, but we also need to ensure that $\|\phi_0\| \leq \|\phi\|$. That is, we need

$$|\phi_G(x) + a\alpha_0| \leq \|\phi\| \|x + a x_0\| \qquad (x \in G, a \in \mathbb{R}).$$

For $x, y \in G$, we have that

$$\phi_G(x) - \phi_G(y) = \phi_G(x - y) \leq \|\phi\| \|x - y\| \leq \|\phi\| \big(\|x + x_0\| + \|y + x_0\| \big).$$

Consequently,

$$-\phi_G(y) - \|\phi\| \|y + x_0\| \leq -\phi_G(x) + \|\phi\| \|x + x_0\|.$$

As x and y were arbitrary,

$$\sup_{y\in G} -\phi_G(y) - \|\phi\|\|y+x_0\| \leq \inf_{x\in G} -\phi_G(x) + \|\phi\|\|x+x_0\|.$$

Hence we can choose α_0 between the inf and the sup. Hence in particular,

$$-\phi_G(x) - \|\phi\|\|x+x_0\| \leq \alpha_0 \leq -\phi_G(x) + \|\phi\|\|x+x_0\| \qquad (x \in G).$$

Re-arranging, we get

$$\big|\alpha_0 + \phi_G(x)\big| \leq \|\phi\|\|x+x_0\|,$$

and so for non-zero $a \in \mathbb{R}$,

$$\big|a\alpha_0 + \phi_G(x)\big| = |a|\big|\alpha_0 + \phi_G(a^{-1}x)\big| \leq |a|\|\phi\|\|a^{-1}x + x_0\| = \|\phi\|\|x + ax_0\|,$$

which shows that $\|\phi_0\| \leq \|\phi_G\|$, as required.

The complex case follows by "complexification". □

The Hahn-Banach theorem tells us that a functional from a subspace can be extended to the whole space without increasing the norm. In particular, extending a functional on a one-dimensional subspace yields the following.

Corollary: Let E be a normed vector, and let $x \in E$. Then there exists $\phi \in E^*$ with $\|\phi\| = 1$ and $\phi(x) = \|x\|$. □

Another useful result which can be proved by Hahn-Banach is the following.

Corollary: Let E be a normed vector, and let F be a subspace of E. For $x \in E$, the following are equivalent:

1. $x \in \overline{F}$ the closure of F;
2. for each $\phi \in E^*$ with $\phi(y) = 0$ for each $y \in F$, we have that $\phi(x) = 0$.

Proof: (1)$\Rightarrow$(2) follows because we can find a sequence (y_n) in F with $y_n \to x$; then it's immediate that $\phi(x) = 0$, because ϕ is continuous. Conversely, we show that if (1) doesn't hold then (2) doesn't hold (that is, the contrapositive to (2)$\Rightarrow$(1)).

So, $x \notin \overline{F}$. Define $\psi : \operatorname{lin}\{F, x\} \to \mathbb{K}$ by

$$\psi(y + tx) = t \qquad (y \in F, t \in \mathbb{K}).$$

This is well-defined, for if $y+tx = y'+t'x$ then either $t = t'$, or otherwise $x = (t-t')^{-1}(y'-y) \in F$ which is a contradiction. The map ψ is obviously linear, so we need to show that it is bounded. Towards a contradiction, suppose that ψ is not bounded, so we can find a sequence $(y_n + t_n x)$ with $\|y_n + t_n x\| \leq 1$ for each n, and yet $|\psi(y_n + t_n x)| = |t_n| \to \infty$. Then $\|t_n^{-1}y_n + x\| \leq 1/|t_n| \to 0$, so that the sequence $(-t_n^{-1}y_n)$, which is in F, converges to x. So x is in the closure of F, a contradiction. So ψ is bounded. By Hahn-Banach, we can find some $\phi \in E^*$ extending ψ. For $y \in F$, we have $\phi(y) = \psi(y) = 0$, while $\phi(x) = \psi(x) = 1$, so (2) doesn't hold, as required. □

$C(X)$ spaces[1]

All our topological spaces are assumed *Hausdorff*. Let X be a compact space, and let $C_{\mathbb{K}}(X)$ be the space of continuous functions from X to $\mathbb{K}$, with pointwise operations, so that $C_{\mathbb{K}}(X)$ is a vector space. We norm $C_{\mathbb{K}}(X)$ by setting

$$\|f\|_\infty = \sup_{x\in X} |f(x)| \qquad (f \in C_{\mathbb{K}}(X)).$$

Theorem: Let X be a compact space. Then $C_{\mathbb{K}}(X)$ is a Banach space. □

[1]This section is not examinable, *BUT* a pre-requisite for the course is the Topology course, and so I assume that you know the basics about compact spaces, and so forth. Standard facts about topology will be used in later sections of the course.

Let E be a vector space, and let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be norms on E. These norms are *equivalent* if there exists $m > 0$ with

$$m^{-1}\|x\|_{(2)} \leq \|x\|_{(1)} \leq m\|x\|_{(2)} \qquad (x \in E).$$

Theorem: Let E be a finite-dimensional vector space with basis $\{e_1, \ldots, e_n\}$, so we can identify E with $\mathbb{K}^n$ as vector spaces, and hence talk about the norm $\|\cdot\|_2$ on E. If $\|\cdot\|$ is any norm on E, then $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent. $\square$

Corollary: Let E be a finite-dimensional normed space. Then a subset $X \subseteq E$ is compact if and only if it is closed and bounded. $\square$

Lemma: Let E be a normed vector space, and let F be a closed subspace of E with $E \neq F$. For $0 < \theta < 1$, we can find $x_0 \in E$ with $\|x_0\| \leq 1$ and $\|x_0 - y\| > \theta$ for $y \in F$. $\square$

Theorem: Let E be an infinite-dimensional normed vector space. Then the *closed unit ball* of E, the set $\{x \in E : \|x\| \leq 1\}$, is not compact.

Proof: Use the above lemma to construct a sequence (x_n) in the closed unit ball of E with, say, $\|x_n - x_m\| \geq 1/2$ for each $n \neq m$. Then (x_n) can have no convergent subsequence, and so the closed unit ball cannot be compact. $\square$

Linear Analysis II: Measure Theory

Updated: 20th October 2009

Let X be a set. A *σ-algebra* on X is a collection of subsets of X, say $\mathcal{R} \subseteq 2^X$, such that $\emptyset, X \in \mathcal{R}$, if $A, B \in \mathcal{R}$, then $A \setminus B \in \mathcal{R}$, and finally, if (A_n) is any sequence in $\mathcal{R}$, then $\bigcup_n A_n \in \mathcal{R}$. For a σ-algebra $\mathcal{R}$ and $A, B \in \mathcal{R}$, we have

$$A \cap B = X \setminus \big(X \setminus (A \cap B)\big) = X \setminus \big((X \setminus A) \cup (X \setminus B)\big) \in \mathcal{R};$$

similarly, $\mathcal{R}$ is closed under taking (countably) infinite intersections.

As the intersection of a family of σ-algebras is again a σ-algebra, and the power set 2^X is a σ-algebra, it follows that given any collection $\mathcal{D} \subseteq 2^X$, there is a σ-algebra $\mathcal{R}$ such that $\mathcal{D} \subseteq \mathcal{R}$, such that if $\mathcal{S}$ is any other σ-algebra, with $\mathcal{D} \subseteq \mathcal{S}$, then $\mathcal{R} \subseteq \mathcal{S}$. We call $\mathcal{R}$ the *σ-algebra generated by $\mathcal{D}$*.

We introduce the symbols $+\infty, -\infty$, and treat these as being "extended real numbers", so $-\infty < t < \infty$ for $t \in \mathbb{R}$. We define $t + \infty = \infty$, $t\infty = \infty$ if $t > 0$ and so forth. We do not (and cannot, in a consistent manner) define $\infty - \infty$ or 0∞.

A *measure* if a map $\mu : \mathcal{R} \to [0, \infty]$ defined on a σ-algebra $\mathcal{R}$, such that $\mu(\emptyset) = 0$, and if (A_n) is a sequence in $\mathcal{R}$ which is pairwise disjoint (that is, $A_n \cap A_m = \emptyset$ for $n \neq m$), then $\mu\big(\bigcup_n A_n\big) = \sum_n \mu(A_n)$. This last property says that μ is *countably additive.* If the sum diverges, then as it will be the sum of positive numbers, we can, without problem, define it to be $+\infty$.

Proposition: Let μ be a measure on a σ-algebra $\mathcal{R}$. Then:

1. If $A, B \in \mathcal{R}$ with $A \subseteq B$, then $\mu(A) \leq \mu(B)$;
2. If $A, B \in \mathcal{R}$ with $A \subseteq B$ and $\mu(B) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$;
3. If (A_n) is a sequence in $\mathcal{R}$, with $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. Then
$$\lim_{n\to\infty} \mu(A_n) = \mu\Big(\bigcup A_n\Big).$$
4. If (A_n) is a sequence in $\mathcal{R}$, with $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. If $\mu(A_m) < \infty$ for some m, then
$$\lim_{n\to\infty} \mu(A_n) = \mu\Big(\bigcap A_n\Big).$$

Measures are useful, but hard to define. An *outer measure* on a set X is a map $\mu^* : 2^X \to [0, \infty]$ such that $\mu^*(\emptyset) = 0$, if $A \subseteq B$ then $\mu^*(A) \leq \mu^*(B)$, and if (A_n) is any sequence in 2^X, then $\mu^*\big(\bigcup_n A_n\big) \leq \sum_n \mu^*(A_n)$. The final condition says that an outer measure is *countably sub-additive.*

The *Lebesgue outer measure* on $\mathbb{R}$ is defined, for $A \subseteq \mathbb{R}$, as

$$\mu^*(A) = \inf\Big\{ \sum_{j=1}^{\infty} (b_j - a_j) : A \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j) \Big\}.$$

We make this definition, as intuitively, the "length", or measure, of the interval (a, b) is $(b - a)$. We can check that μ^* *is* an outer measure.

For example, $\mu^*(A) = 0$ for any countable set, which follows, as clearly $\mu^*(\{x\}) = 0$ for any $x \in \mathbb{R}$.

Lemma: Let $a < b$. Then $\mu^*([a,b]) = b - a$.

Proof: For $\epsilon > 0$, as $[a,b] \subseteq (a-\epsilon, b+\epsilon)$, we have that $\mu^*([a,b]) \leq b - a + 2\epsilon$. As $\epsilon > 0$, was arbitrary, $\mu^*([a,b]) \leq b - a$.

Let $[a,b] \subseteq \bigcup_n (a_n, b_n)$, where we assume that $a_n < b_n$ for each n. As $[a,b]$ is compact, by re-ordering, there exists some N with $[a,b] \subseteq \bigcup_{n=1}^N (a_n, b_n)$. Again, by re-ordering, we may suppose that $a_1 \leq a_2 \leq \cdots \leq a_N$. If $b_k \geq b_j$ for some $k < j$, then $(a_j, b_j) \subseteq (a_k, b_k)$, and so we can remove (a_j, b_j). Hence we may suppose that $b_1 < b_2 < \cdots < b_N$. If $b_1 \leq a$, then (a_1, b_1) does not cover any of $[a,b]$, so we can remove (a_1, b_1). So we may suppose that $a_1 < a < b_1$, and similarly, that $b_{N-1} \leq b$. Suppose, towards a contradiction, that $\sum_{n=1}^N (b_n - a_n) < b - a$. If $a_{k+1} < b_k$ for $1 \leq k < N$, then

$$\sum_{n=1}^N b_n - a_n > \sum_{n=1}^{N-1} a_{n+1} - a_n + b_N - a_N = b_N - a_1 > b - a_1 > b - a,$$

a contradiction. Hence $a_{k+1} \geq b_k$ for some k. As $a < b_1 \leq b_k \leq b_{N-1} \leq b$, we have that $b_k \in [a,b]$, and so $b_k \in (a_j, b_j)$ for some j. As $b_1 < b_2 < \cdots < b_{k-1} < b_k$, we must have that $j > k$. However, then $b_k \leq a_{k+1} \leq a_j$, a contradiction. We conclude that $\sum_{n=1}^N b_n - a_n \geq b - a$, and hence that $\mu^*([a,b]) = b - a$. $\square$

Our next aim is to construct measures from outer measures. Given an outer measure μ^*, we define $E \subseteq X$ to be *measurable* if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E) \qquad (A \subseteq X).$$

As μ^* is sub-additive, this is equivalent to

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E) \qquad (A \subseteq X),$$

as the other inequality is automatic.

Theorem: Let μ^* be an outer measure on X, and let $\mathcal{R}$ be the collection of all measurable sets for μ^*. Then $\mathcal{R}$ is a σ-algebra, and if μ is the restriction of μ^* to $\mathcal{R}$, then μ is a measure.

This is useful, but notice that $\{X, \emptyset\}$ is a σ-algebra, so we have to check that the σ-algebra this theorem constructs is usefully large. This is the case for the Lebesgue outer measure.

Proposition: For any $x \in \mathbb{R}$, the sets $(-\infty, x]$ and $[x, \infty)$ are Lebesgue measurable (that is, measurable with respect to the Lebesgue outer measure).

Proof: The idea is to take $A \subseteq \mathbb{R}$, and for an open cover of A, we can split up the intervals in the open cover to give covers for $(-\infty, x] \cap A$ and $(x, \infty) \cap A$ (some care is required, as $(-\infty, x]$ is closed, so you have to tweak the open intervals). $\square$

Corollary: Let $E \subseteq \mathbb{R}$ be open or closed. Then E is Lebesgue measurable.

Proof: This is a common trick, using the density and the countability of the rationals. As σ-algebras are closed under taking complements, we need only show that open sets are Lebesgue measurable. As $(-\infty, x) = \mathbb{R} \setminus [x, \infty)$, it is Lebesgue measurable, as is (x, ∞), for each $x \in \mathbb{R}$. Hence $(a,b) = (-\infty, b) \cap (a, \infty)$ is also Lebesgue measurable, for any $a < b$.

Now let $U \subseteq \mathbb{R}$ be open. For each $x \in U$, there exists $a < b$ with $x \in (a,b) \subseteq \mathbb{R}$. By making a slightly larger, and b slightly smaller, we can ensure that $a, b \in \mathbb{Q}$. Let $\mathcal{D} = \{(a,b) : a, b \in \mathbb{Q}, a < b\}$, and let $\mathcal{D}_U = \{W \in \mathcal{D} : W \subseteq U\}$. Thus $U = \bigcup \mathcal{D}_U$. Notice that each member of $\mathcal{D}$ is Lebesgue measurable, and $\mathcal{D}$ is countable. The same facts are true for $\mathcal{D}_U$, and thus U is the countable (or finite) union of Lebesgue measurable sets, and hence U is Lebesgue measurable itself. $\square$

We call a measure μ defined on $\mathcal{R}$ *complete* if whenever $E \subseteq X$ is such that there exists $F \in \mathcal{R}$ with $\mu(F) = 0$ and $E \subseteq F$, we have that $E \in \mathcal{R}$. Measures constructed from outer measures by the above theorem are always complete. On the example sheet, we saw how to form a complete measure from a given measure. We call sets like E *null sets*: complete measures are useful, because it is useful to be able to say that null sets are in our σ-algebra. Null sets can be quite complicated. For the Lebesgue measure, all countable subsets of $\mathbb{R}$ are null, but then so is the Cantor set, which is uncountable.

Integration

We now come to the main use of measure theory: to define a general theory of integration. From now on, by a *measure space* we shall mean a triple $(X, \mathcal{R}, \mu)$, where X is a set, $\mathcal{R}$ is a σ-algebra on X, and μ is a measure defined on $\mathcal{R}$. We say that the members of $\mathcal{R}$ are *measurable*, or $\mathcal{R}$-measurable, if necessary to avoid confusion.

A function $f : X \to \mathbb{K}$ is *measurable* if $f^{-1}(U) \in \mathcal{R}$ for each open $U \subseteq \mathbb{K}$. We shall mostly work with $\mathbb{K} = \mathbb{R}$ in what follows: the complex numbers will be used in later chapters.

Lemma: Let $f, g : X \to \mathbb{R}$ be measurable. Then $f + g, fg, \max(f, g)$ and $\min(f, g)$ are all measurable.

Proof: See the example sheet. We shall repeatedly use these results in what follows. $\square$

A function $f : X \to \mathbb{R}$ is *simple* if there exist $A_1, \cdots, A_n \in \mathcal{R}$ which are pairwise disjoint, and there exist $t_1, \cdots, t_n \in \mathbb{R}$, such that

$$f(x) = \begin{cases} t_k & : x \in A_k \text{ for some } 1 \le k \le n, \\ 0 & : \text{otherwise.} \end{cases}$$

For $A \subseteq X$, we define χ_A to be the *indicator function* of A, by

$$\chi_A(x) = \begin{cases} 1 & : x \in A, \\ 0 & : x \notin A. \end{cases}$$

Then, if χ_A is measurable, then $\chi_A^{-1}((1/2, 3/2)) = A \in \mathcal{R}$; conversely, if $A \in \mathcal{R}$, then $X \setminus A \in \mathcal{R}$, and we see that for any $U \subseteq \mathbb{R}$ open, $\chi_A^{-1}(U)$ is either $\emptyset$, A, $X \setminus A$, or X, all of which are in $\mathcal{R}$. So χ_A is measurable if and only if $A \in \mathcal{R}$.

Lemma: A function $f : X \to \mathbb{R}$ is simple if and only if

$$f = \sum_{k=1}^{n} t_k \chi_{A_k}$$

for some $(t_k)_{k=1}^n \subseteq \mathbb{R}$ and $A_1, \cdots, A_k \in \mathcal{R}$. That is, simple functions are linear combinations of indicator functions of measurable sets.

Proof: The easiest way to prove this is to first prove that a function is simple if and only if its image is a finite subset of $\mathbb{R}$. Notice that it is now obvious that the collection of simple functions forms a vector space: this wasn't clear from the original definition. $\square$

We define the *integral* of a simple function $f : X \to [0, \infty)$ by setting

$$\int f \, d\mu = \int_X f \, d\mu = \sum_{k=1}^{n} t_k \mu(A_k) \quad \text{if } f \text{ has the representation} \quad f = \sum_{k=1}^{n} t_k \chi_{A_k}.$$

In this special setting, we define $0\mu(A) = 0$ for any A, even if $\mu(A) = \infty$. We allow the integral to be ∞. It is another combinatorial exercise to show that this definition is independent of the way we write f.

Then, for an arbitrary measurable $f : X \to [0,\infty)$, we define

$$\int_X f \; d\mu = \sup\Big\{ \int_X g \; d\mu : g \geq 0, g \text{ is simple}, g \leq f \Big\},$$

where the supremum may be infinite. When it's finite, we say that f is *integrable*. Notice that by "$g \leq f$", I mean that $g(x) \leq f(x)$ for all $x \in X$.

Finally, for a measurable $f : X \to \mathbb{R}$, we say that f is *integrable* if $|f|$ is integrable. Notice that $|f| = f_+ + f_-$ where

$$f_+ = \max(f, 0), \qquad f_- = -\min(f, 0),$$

which are measurable functions, and so $|f|$ is also measurable. Furthermore, $|f|$ is integrable if and only if f_+ and f_- are both integrable, and we define

$$\int_X f \; d\mu = \int_X f_+ \; d\mu - \int_X f_- \; d\mu.$$

Showing that integration satisfies even simple properties requires a bit of a detour.

Proposition: Let $f, g : X \to [0,\infty)$ be simple, and let $a, b \in [0,\infty)$. Then:

1. $\int_X af + bg \; d\mu = a \int_X f \; d\mu + b \int_X g \; d\mu$;
2. If $f \leq g$ then $\int_X f \; d\mu \leq \int_X g \; d\mu$;
3. $\int_X f \; d\mu = 0$ if and only if $\mu(\{x \in X : f(x) \neq 0\}) = 0$.

Proof: This is another slightly tedious combinatorial exercise. If you can prove that the integral of a simple function is well-defined, in the sense that it is independent of the way we choose to write the simple function, then the rest of the proposition is easy.[1] $\square$

The following important theorem allows us to draw conclusions about the *pointwise* limit of functions (which is, say, a rare thing when dealing with continuous functions: you'd normally need some sense of uniform convergence).

Theorem (Monotone Convergence): Let $f_n : X \to [0,\infty)$ be a sequence of measurable functions, and let $f : X \to [0,\infty)$ be a function. Suppose that for each $x \in X$, we have $f_1(x) \leq f_2(x) \leq \cdots$, and $f(x) = \lim_n f_n(x)$. Then f is measurable, and

$$\int_X f \; d\mu = \lim_{n\to\infty} \int_X f_n \; d\mu.$$

Proof: We give the proof, as it illustrates the general strategy for many proofs in this area. Firstly, we show that f is measurable. By the arguments we used above, it is enough to show

[1] In previous years, I have stated this result for functions mapping into $\mathbb{R}$, not $[0,\infty)$. I also stated that a simple function $f : X \to \mathbb{R}$ is integrable if and only if f admits a representation

$$f = \sum_{k=1}^n t_k \chi_{A_k}$$

where $(t_k)_{k=1}^n \subseteq \mathbb{R}$ and $A_1, \cdots, A_k$ are measurable and pairwise disjoint, with

$$\sum_{k=1}^n |t_k| \mu(A_k) < \infty.$$

I think you could prove these results fairly easily as well, by splitting f into f_+ and f_-, but it seemed easier to work with positive functions, and then deduce the general case later.

that $f^{-1}((-\infty,t]), f^{-1}([t,\infty)) \in \mathcal{R}$ for each $t \in \mathbb{R}$. Notice that

$$\begin{aligned} x \in f^{-1}((-\infty,t]) &\Leftrightarrow \lim_n f_n(x) = f(x) \leq t \\ &\Leftrightarrow f_n(x) \leq t \qquad \text{for all } n \\ &\Leftrightarrow x \in \bigcap_{n=1}^{\infty} f_n^{-1}\big((-\infty,t]\big). \end{aligned}$$

However, for each n, $f_n^{-1}((-\infty,t]) \in \mathcal{R}$ as f_n is measurable, and so $f^{-1}((-\infty,t]) \in \mathcal{R}$. Now notice that

$$\begin{aligned} x \in f^{-1}([t,\infty)) &\Leftrightarrow \lim_n f_n(x) = f(x) \geq t \\ &\Leftrightarrow \lim_n f_n(x) > t-\epsilon \qquad \text{for all } \epsilon > 0 \\ &\Leftrightarrow \forall\, \epsilon > 0\ \exists\, n \geq 1,\ f_n(x) > t - \epsilon \\ &\Leftrightarrow x \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} f_n^{-1}((t-1/m,\infty)). \end{aligned}$$

This is more complicated, but we can still conclude that $f^{-1}([t,\infty)) \in \mathcal{R}$.

I. Suppose $f_n = \chi_{A_n}$ for each n, so we must have that $A_n \in \mathcal{R}$ for all n, and $A_1 \subseteq A_2 \subseteq \cdots$ as the f_n are increasing. Similarly, then $f = \chi_A$ where $A = \bigcup_n A_n$, and so the result follows from a proposition above, as

$$\lim_n \int_X f_n \ d\mu = \lim_n \mu(A_n) = \mu\Big(\bigcup_n A_n\Big) = \mu(A) = \int_X f \ d\mu.$$

II. Suppose that $f = \chi_A$, and each f_n is simple. For $\epsilon > 0$, if we set $A_n = \{x \in X : f_n(x) > 1-\epsilon\} = f_n^{-1}((1-\epsilon,\infty)) \in \mathcal{R}$, then $A_1 \subseteq A_2 \subseteq \cdots$ and, assuming $\epsilon < 1$, we have that

$$x \in A \ \Leftrightarrow\ f(x) > 1-\epsilon \ \Leftrightarrow\ \lim_n f_n(x) > 1-\epsilon \ \Leftrightarrow\ \exists\, n,\ x \in A_n.$$

So $A = \bigcup_n A_n$. Also notice that for each n,

$$(1-\epsilon)\chi_{A_n} \leq f_n \leq f = \chi_A.$$

As each f_n is simple, by the proposition above, as $f_1 \leq f_2 \leq \cdots$ we have that $\int_X f_1 \ d\mu \leq \int_X f_2 \ d\mu \leq \cdots$. Also, as f is simple and $f_n \leq f$, we have that $\int_X f_n \ d\mu \leq \int_X f \ d\mu$ for each n. Hence $\lim_{n\to\infty} \int_X f_n \ d\mu$ exists.

So, in conclusion,

$$(1-\epsilon)\mu(A_n) \leq \int_X f_n \ d\mu \leq \int_X f \ d\mu = \mu(A).$$

By part 1, we know that $\mu(A_n) \to \mu(A)$, so as $\epsilon > 0$ was arbitrary, and the sandwich rule, we must have that $\lim_{n\to\infty} \int_X f_n \ d\mu = \int_X f \ d\mu$ as required.

III. Suppose that each f_n is simple, and that f is simple. Let

$$f = \sum_{k=1}^{N} t_k \chi_{A_k},$$

where we may suppose that $t_k > 0$ for each k, and the (A_k) are pairwise disjoint. As f is measurable, we see that $A_k \in \mathcal{R}$ for each k. Then we see that $t_k^{-1}\chi_{A_k} f_n \uparrow t_k^{-1}\chi_{A_k} f = \chi_{A_k}$, where $\uparrow$ means "increasing pointwise". So by part 2, we have that

$$\lim_n \int_X t_k^{-1}\chi_{A_k} f_n \ d\mu = \mu(A_k).$$

As the (A_k) are pairwise disjoint, and $f(x) = 0$ if $x \notin A_1 \cup \cdots \cup A_N$, we must have that

$$f_n = \sum_{k=1}^{n} f_n \chi_{A_k}$$

for all n. Hence

$$\lim_n \int_X f_n \, d\mu = \lim_n \sum_{k=1}^{N} \int_X f_n \chi_{A_k} \, d\mu = \sum_{k=1}^{N} t_k \mu(A_k) = \int_X f \, d\mu.$$

IV. Now suppose that each f_n is simple, and that $f \geq 0$ is measurable. Let g be simple with $0 \leq g \leq f$. Then $\min(g, f_n)$ is simple for each n, and $\min(g, f_n) \uparrow g$ as $f_n \uparrow f$ and $g \leq f$. So by part 3,

$$\lim_n \int_X \min(g, f_n) \, d\mu = \int_X g \, d\mu.$$

As $\min(g, f_n) \leq f_n$ for each n, we have

$$\int_X g \, d\mu = \lim_n \int_X \min(g, f_n) \, d\mu \leq \lim_n \int_X f_n \, d\mu.$$

Thus, by the definition of the integral,

$$\int_X f \, d\mu \leq \lim_n \int_X f_n \, d\mu,$$

as $\int_X f \, d\mu$ is the supremum of $\int_X g \, d\mu$. However, each f_n is simple and $f_n \leq f$, so by the definition of the integral, certainly

$$\int_X f_n \, d\mu \leq \int_X f \, d\mu \ (\forall\, n) \implies \lim_n \int_X f_n \, d\mu \leq \int_X f \, d\mu,$$

so we must have equality.

V. We finally do the general case. For each $t \in \mathbb{R}$, let $\lfloor t \rfloor$ be the largest integer which is less than or equal to t. Then define, for each n,

$$g_n = \min\big(n, 2^{-n} \lfloor 2^n f_n \rfloor\big).$$

A moment's thought should reveal that each g_n is simple, and that $g_n \leq f_n$. Thus

$$\int_X g_n \, d\mu \leq \int_X f_n \, d\mu \qquad (n \geq 1).$$

If h is a simple function with $h \leq f_n$ then as $f_n \leq f$, also $h \leq f$. Taking the supremum over all such h shows that

$$\int_X f_n \, d\mu \leq \int_X f \, d\mu \qquad (n \geq 1).$$

For $x \in X$, if n is much larger than $f(x)$, then as $\lfloor 2^n f_n(x) \rfloor \leq 2^n f_n(x) < \lfloor 2^n f_n(x) \rfloor + 1$, we have that

$$f_n(x) - 2^{-n} < g_n(x) = 2^{-n} \lfloor 2^n f_n(x) \rfloor \leq f_n(x) \leq f(x).$$

Thus $f_n \uparrow f$ implies that also $g_n \uparrow f$, and so by part 4,

$$\int_X f \, d\mu = \lim_n \int_X g_n \, d\mu \leq \lim_n \int_X f_n \, d\mu \leq \int_X f \, d\mu,$$

so we have equality throughout, as required. $\square$

Let us pull some *important* things out of this proof. The notation $f_n \uparrow f$ means that for each $x \in X$, both $f_1(x) \leq f_2(x) \leq \cdots$ and $f_n(x) \to f(x)$. Given a measurable, positive f, letting

$$f_n = \min\left(n, 2^{-n}\lfloor 2^n f_n \rfloor\right)$$

is a standard, and very useful, way of getting some simple functions[2] with $f_n \uparrow f$.

Theorem: Let $f, g : X \to [0, \infty)$ be measurable, and let $a, b \geq 0$. Then:

1. $\int_X af + bg \; d\mu = a \int_X f \; d\mu + b \int_X g \; d\mu$;
2. If $f \leq g$, then $\int_X f \; d\mu \leq \int_X g \; d\mu$;
3. $\int_X f \; d\mu = 0$ if and only if $\mu(\{x \in X : f(x) \neq 0\}) = 0$.

Proof: We show (1). Let $f_n = \min(n, 2^{-n}\lfloor 2^n f \rfloor)$ and $g_n = \min(n, 2^{-n}\lfloor 2^n g \rfloor)$. Hence f_n and g_n are simple functions, for each n, and hence

$$\int_X af_n + bg_n \; d\mu = a \int_X f_n \; d\mu + b \int_X f_n \; d\mu.$$

We can check that $f_1(x) \leq f_2(x) \leq \cdots$ and $f(x) = \lim_n f_n(x)$, for each $x \in X$. We write $f_n \uparrow f$ to denote this. As $g_n \uparrow g$ and also $af_n + bg_n \uparrow af + bg$, by Monotone Convergence,

$$\begin{aligned} \int_X af + bg \; d\mu &= \lim_n \int_X af_n + bg_n \; d\mu = \lim_n a \int_X f_n \; d\mu + b \int_X f_n \; d\mu \\ &= a \int_X f \; d\mu + b \int_X g \; d\mu, \end{aligned}$$

as required. It is worth noting that this result seems impossible to prove without using Monotone Convergence, or some similar tool. $\square$

Theorem: Let $f, g : X \to \mathbb{R}$ be measurable, and let $a, b \in \mathbb{R}$. Then:

1. $\int_X af + bg \; d\mu = a \int_X f \; d\mu + b \int_X g \; d\mu$;
2. If $f \leq g$, then $\int_X f \; d\mu \leq \int_X g \; d\mu$;
3. If $\mu(\{x \in X : f(x) \neq 0\}) = 0$, then $\int_X f \; d\mu = 0$.

Proof: We show (2). Notice that $f_+ - f_- = f \leq g = g_+ - g_-$, so by rearranging, $f_+ + g_- \leq g_+ + f_-$, and so by the previous theorem, $\int_X f_+ + g_- \; d\mu \leq \int_X g_+ + f_- \; d\mu$, as all the functions are positive. By splitting up the integrals and rearranging again, and then recombining the integrals, we get $\int_X f \; d\mu \leq \int_X g \; d\mu$, as required. $\square$

We now recall the notion of the Riemann integral. Instead of defining it here, we shall simply remember the Fundamental Theorem of Calculus. Namely, if $f : [a, b] \to \mathbb{R}$ is continuous, and we define $F : [a, b] \to \mathbb{R}$ by

$$F(t) = \int_a^t f(x) \; dx \qquad (a \leq t \leq b),$$

then F is differentiable on (a, b), and $F'(t) = f(t)$ for $a < t < b$.

Theorem: Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is Lebesgue integrable, and if we define $G : [a, b] \to \mathbb{R}$ by

$$G(t) = \int_{\mathbb{R}} f \chi_{[a,t]} \; d\mu,$$

[2] It's unimportant that we use n and 2^n: any increasing sequences would work.

then G is differentiable on (a,b), and $G'(t) = f(t)$ for $a < t < b$. As a result, the Riemann and Lebesgue integrals of f over $[a,b]$ agree.

Proof: As f is continuous on a compact set, it is bounded, and hence $|f|$ is bounded, and so has finite integral (as also $[a,b]$ has finite measure). Hence f is Lebesgue integrable. We show that G is differentiable in a similar way to the proof of the Fundamental Theorem of Calculus for the Riemann integral.

Then, as $G' = F'$ on (a,b), and $G(a) = F(a)$, we conclude that $G(b) = F(b)$, that is,

$$\int_a^b f(x)\ dx = \int_{\mathbb{R}} f\chi_{[a,b]}\ d\mu = \int_{[a,b]} f\ d\mu.$$

From now on, we shall hence identify the Riemann Integral and the Lebesgue Integral, at least for continuous functions defined on closed intervals. □

If P is a property of the points of a measure space $(X, \mathcal{R}, \mu)$, then we say that P holds *almost everywhere* if $\mu(\{x \in X : P(x) \text{ not true }\}) = 0$. For example, if $f, g : X \to \mathbb{K}$ are measurable functions, then we say that $f = g$ almost everywhere (or a.e.) if $\mu(\{x \in X : f(x) \neq g(x)\}) = 0$.

Theorem (Fatou's Lemma): Let (f_n) be a sequence of measurable functions $X \to [0,\infty)$, and define

$$f(x) = \begin{cases} \liminf_n f_n(x) & : \text{if } \liminf_n f_n(x) < \infty, \\ 0 & : \text{otherwise.} \end{cases}$$

Then f is measurable, and

$$\int_X f\ d\mu \leq \liminf_n \int_X f_n\ d\mu.$$

This version of Fatou's Lemma might be different to the statement you can find in a book[3]. This is because I have decided that I do not want to allow functions to take the value ∞ (or $-\infty$).

Theorem (Dominated Convergence Theorem): Let (f_n) be a sequence of measurable functions $X \to \mathbb{R}$ such that $f(x) = \lim_{n\to\infty} f_n(x)$ exists for almost every $x \in X$; define $f(x) = 0$ otherwise. Suppose furthermore that for some integrable $g : X \to [0,\infty)$, we have that $|f_n| \leq g$ almost everywhere, for each n. Then f is integrable, and

$$\int_X f\ d\mu = \lim_{n\to\infty} \int_X f_n\ d\mu.$$

In words, this means that we can push a pointwise limit through the integral, as long as every function in question is dominated by some functions with finite integral. The Monotone Convergence Theorem tells us the same, under the condition that the limit is increasing. We have seen examples which show that we cannot push a pointwise limit through an integral without one of these extra conditions.

Notation

For a measure space $(X, \mathcal{R}, \mu)$ and an integrable function $f : X \to \mathbb{R}$, we have

$$\int_X f\ d\mu = \int f\ d\mu = \int_X f(x)\ d\mu(x).$$

The form of the right hand side is useful when, say, f might depend upon two variables (for example, as in the next section).

[3] Or on wikipedia

When $Y \subseteq X$ is measurable (that is, $Y \in \mathcal{R}$) we saw on the example sheet that

$$\mathcal{R}_Y := \{A \cap Y : A \in \mathcal{R}\} = \{A \in \mathcal{R} : A \subseteq Y\}$$

is a σ-algebra on Y. Furthermore, the restriction of μ to $\mathcal{R}_Y$ defines a measure on $\mathcal{R}_Y$, so $(Y, \mathcal{R}_Y, \mu)$ becomes a measure space.

In particular, when $[a, b]$ is a closed interval of $\mathbb{R}$, we can restrict the Lebesgue measure to $[a, b]$. Clearly every open subset of $[a, b]$ is then Lebesgue measurable. If $f : [a, b] \to \mathbb{R}$ is integrable, then we can either integrate it over the restriction of Lebesgue measure on $[a, b]$; or we can multiply f by $\chi_{[a,b]}$ and regard $f\chi_{[a,b]}$ as a function $\mathbb{R} \to \mathbb{R}$, and then integrate this. A quick check shows that we get the same answer, and so we write

$$\int_{[a,b]} f \; d\mu = \int_{\mathbb{R}} f\chi_{[a,b]} \; d\mu.$$

We try to *avoid* the notation $\int_a^b f \; d\mu$, as this could be confused with Riemann integration. Many books will use this, however.

Product Measures

How do we handle integration on $\mathbb{R}^2$? We could re-develop Lebesgue measure, say using rectangles instead of intervals as our "test" sets. This is boring, and would also make doing calculations hard. From Multi-variate Calculus, we expect to be able to split up integration over $\mathbb{R}^2$ into two integrals over $\mathbb{R}$.

Formally, we wish to handle $X \times Y$ for measure spaces X and Y. To simplify things, we shall mostly work with *finite* measure spaces, that is, when the measure of the whole space is finite. So $[0, 1]$ with Lebesgue measure is finite, but $\mathbb{R}$ is not. The results do hold for spaces like $\mathbb{R}$, however (but the proofs become even more technical).

Let X and Y be spaces, and let $\mathcal{R}$ and $\mathcal{S}$ be σ-algebras on X and Y respectively. Then $\mathcal{R} \otimes \mathcal{S}$ is the σ-algebra *generated* by $\{A \times B : A \in \mathcal{R}, B \in \mathcal{S}\}$. These sets are sort of "generalised rectangles", but $\mathcal{R} \otimes \mathcal{S}$ contains everything we can get by taking countably infinite unions, set differences, and so forth.[4]

Lemma: Let $f : X \times Y \to \mathbb{R}$ be measurable with respect to $\mathcal{R} \otimes \mathcal{S}$. For each $x \in X$, let $f_x : Y \to \mathbb{R}$ be the slice of f, defined by $f_x(y) = f(x, y)$. Then f_x is $\mathcal{S}$-measurable.

Lemma: Suppose that $(Y, \mathcal{S}, \lambda)$ is a finite measure space. Let $f : X \times Y \to \mathbb{R}$ be $\mathcal{R} \otimes \mathcal{S}$-measurable, and bounded. Then the function $f_1 : X \to \mathbb{R}$, defined by

$$f_1(x) = \int_Y f_x \; d\lambda = \int_Y f(x, y) \; d\lambda(y), \qquad (x \in X)$$

is bounded and $\mathcal{R}$-measurable.

Theorem: Now suppose that both $(X, \mathcal{R}, \mu)$ and $(Y, \mathcal{S}, \lambda)$ are finite measure spaces. There exists a unique measure $\mu \otimes \lambda$ on $\mathcal{R} \otimes \mathcal{S}$ such that

$$(\mu \otimes \lambda)(A \times B) = \mu(A)\lambda(B) \qquad (A \in \mathcal{R}, B \in \mathcal{S}).$$

Proof: For $A \in \mathcal{R} \otimes \mathcal{S}$, $f = \chi_A$ is measurable and bounded, and so we can apply the above lemmas to see that f_1 is bounded and $\mathcal{R}$-measurable. As X is finite, f_1 is integrable, and so we can define

$$(\mu \otimes \lambda)(A) = \int_X f_1 \; d\mu = \int_X \int_Y \chi_A(x, y) \; d\lambda(y) \; d\mu(x).$$

[4] $\mathcal{R} \otimes \mathcal{S}$ is *very* complicated, which is why the proofs in this section are so complicated.

Monotone convergence needs to be used to show that this is a measure. $\square$

We defined f_1 above when f was assumed bounded. Now suppose that $f : X \times Y \to [0, \infty)$ is $(\mathcal{R} \otimes \mathcal{S})$-measurable. Define

$$f_1(x) = \begin{cases} \int_Y f_x \; d\lambda & \text{when this is finite,} \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\int_X \int_Y f(x, y) \; d\lambda(y) \; d\mu(x) \text{ is defined to be } \int_X f_1 \; d\mu.$$

Proposition: Let $f : X \times Y \to [0, \infty)$ be $(\mathcal{R} \otimes \mathcal{S})$-measurable, and $\mu \otimes \lambda$-integrable. Then

$$\int_{X \times Y} f \; d(\mu \otimes \lambda) = \int_X \int_Y f(x, y) \; d\lambda(y) \; d\mu(x).$$

Proposition: Let $f : X \times Y \to [0, \infty)$ be $(\mathcal{R} \otimes \mathcal{S})$-measurable. Then f is $(\mu \otimes \lambda)$-integrable if and only if:

1. $\int_Y f(x, y) \; d\lambda(y) = \int_Y f_x \; d\lambda < \infty$ almost everywhere, with respect to μ; and
2. $\int_X f_1 \; d\mu < \infty$.

Finally, suppose that $f : X \times Y \to \mathbb{R}$ is $(\mathcal{R} \otimes \mathcal{S})$-measurable. Define

$$f_1(x) = \begin{cases} \int_Y f_x \; d\lambda & \text{when } \int_Y |f_x| \; d\lambda < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\int_Y |f_x| \; d\lambda < \infty$ if and only if f_x is λ-integrable.

Theorem (Fubini's Theorem): Let $f : X \times Y \to \mathbb{R}$ be $(\mathcal{R} \otimes \mathcal{S})$-measurable. If f is $\mu \otimes \lambda$-integrable, then:

1. the map $f_x : Y \to \mathbb{R}; y \mapsto f(x, y)$ is λ-integrable for μ-almost every $x \in X$;
2. the map $f_1 : X \to \mathbb{R}$ is μ-integrable.
3. $$\int_{X \times Y} f \; d(\mu \otimes \lambda) = \int_X \int_Y f(x, y) \; d\lambda(y) \; d\mu(x),$$

   where the right-hand-side is *defined* to be $\int_X f_1 \; d\mu$.

We have done everything on the "right" first and then on the "left" (that is, integrate first over Y, and then over X). However, we could do everything the other way around. This yields a corollary, which is actually how Fubini's Theorem is most often applied.

Corollary: Let $f : X \times Y \to \mathbb{R}$ be $(\mathcal{R} \otimes \mathcal{S})$-measurable. If f is $\mu \otimes \lambda$-integrable, then

$$\int_X \int_Y f(x, y) \; d\lambda(y) \; d\mu(x) = \int_Y \int_X f(x, y) \; d\mu(x) \; d\lambda(y).$$

Deciding if f is $(\mathcal{R} \otimes \mathcal{S})$-measurable seems to be hard. But if, for example, $X = Y = [0, 1]$ with the Lebesgue measure, then if f is continuous (except maybe at finitely many points) then f is certainly $(\mathcal{R} \otimes \mathcal{S})$-measurable. We have seen examples of such f which are not $\mu \otimes \lambda$-integrable, and for which the order of integration does matter.

Linear Analysis III: $\mathcal{L}^p$ spaces

Updated: 15th November 2009

Let $(X, \mathcal{R}, \mu)$ be a measure space. For $1 \leq p < \infty$, we define $L^p(\mu)$ to be the space of measurable functions $f : X \to \mathbb{K}$ such that

$$\int_X |f|^p \, d\mu < \infty.$$

We define $\| \cdot \|_p : L^p(\mu) \to [0, \infty)$ by

$$\|f\|_p = \Big(\int_X |f|^p \, d\mu \Big)^{1/p} \qquad (f \in L^p(\mu)).$$

Notice that if $f = 0$ almost everywhere, then $|f|^p = 0$ almost everywhere, and so $\|f\|_p = 0$. However, there can be non-zero functions such that $f = 0$ almost everywhere. So $\| \cdot \|_p$ is *not* a norm on $L^p(\mu)$.

Lemma: Let $1 < p < \infty$, let $q \in (1, \infty)$ be such that $1/p + 1/q = 1$. For $f \in L^p(\mu)$ and $g \in L^q(\mu)$, we have that fg is integrable, and

$$\int_X |fg| \, d\mu \leq \|f\|_p \|g\|_q.$$

Proof: Recall that we know that

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q} \qquad (a, b \in \mathbb{K}).$$

Define measurable functions $a, b : X \to \mathbb{K}$ by setting

$$a(x) = \frac{f(x)}{\|f\|_p}, \quad b(x) = \frac{g(x)}{\|g\|_q} \qquad (x \in X).$$

So we have that

$$|a(x)b(x)| \leq \frac{|f(x)|^p}{p\|f\|_p^p} + \frac{|g(x)|^q}{q\|g\|_q^q} \qquad (x \in X).$$

By integrating, we see that

$$\int_X |ab| \, d\mu \leq \frac{1}{p\|f\|_p^p} \int_X |f|^p \, d\mu + \frac{1}{q\|g\|_q^q} \int_X |g|^q \, d\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

Hence, by the definition of a and b,

$$\int_X |fg| \leq \|f\|_p \|g\|_q,$$

as required. $\square$

Lemma: Let $f, g \in L^p(\mu)$ and let $a \in \mathbb{K}$. Then:

1. $\|af\|_p = |a| \|f\|_p$;
2. $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

In particular, L^p is a vector space.

Proof: Part (1) is easy. For (2), we need a version of Minkowski's Inequality, which will follow from the previous lemma. Notice that the $p = 1$ case is easy, so suppose that $1 < p < \infty$. We have that

$$\begin{aligned}\int_X |f+g|^p \, d\mu &= \int_X |f+g|^{p-1}|f+g| \, d\mu \\ &\leq \int_X |f+g|^{p-1}\big(|f|+|g|\big) \, d\mu \\ &= \int_X |f+g|^{p-1}|f| \, d\mu + \int_X |f+g|^{p-1}|g| \, d\mu.\end{aligned}$$

Applying the lemma, this is

$$\leq \|f\|_p \Big(\int_X |f+g|^{q(p-1)} \, d\mu \Big)^{1/q} + \|g\|_p \Big(\int_X |f+g|^{q(p-1)} \, d\mu \Big)^{1/q}.$$

As $q(p-1) = p$, we see that

$$\|f+g\|_p^p \leq \big(\|f\|_p + \|g\|_p\big)\|f+g\|_p^{p/q}.$$

As $p - p/q = 1$, we conclude that

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

as required.

In particular, if $f, g \in L^p(\mu)$ then $af + g \in L^p(\mu)$, showing that $L^p(\mu)$ is a vector space. $\square$

We define an equivalence relation $\sim$ on the space of measurable functions by setting $f \sim g$ if and only if $f = g$ almost everywhere. We can check that $\sim$ is an equivalence relation (the slightly non-trivial part is that $\sim$ is transitive).

Proposition: For $1 \leq p < \infty$, the collection of equivalence classes $L^p(\mu)/\sim$ is a vector space, and $\|\cdot\|_p$ is a well-defined norm on $L^p(\mu)/\sim$.

Proof: We need to show that addition, and scalar multiplication, are well-defined on $L^p(\mu)/\sim$. Let $a \in \mathbb{K}$ and $f_1, f_2, g_1, g_2 \in L^p(\mu)$ with $f_1 \sim f_2$ and $g_1 \sim g_2$. Then it's easy to see that $af_1 + g_1 \sim af_2 + g_2$; but this is all that's required!

If $f \sim g$ then $|f|^p = |g|^p$ almost everywhere, and so $\|f\|_p = \|g\|_p$. So $\|\cdot\|_p$ is well-defined on equivalence classes. In particular, if $f \sim 0$ then $\|f\|_p = 0$. Conversely, if $\|f\|_p = 0$ then $\int_X |f|^p \, d\mu = 0$, so as $|f|^p$ is a positive function, we must have that $|f|^p = 0$ almost everywhere. Hence $f = 0$ almost everywhere, so $f \sim 0$. That is,

$$\big\{f \in L^p(\mu) : f \sim 0\big\} = \big\{f \in L^p(\mu) : \|f\|_p = 0\big\}.$$

It follows from the above lemma that this is a subspace of $L^p(\mu)$.

The above lemma now immediately shows that $\|\cdot\|_p$ is a norm on $L^p(\mu)/\sim$. $\square$

We write[1] $\mathcal{L}^p(\mu)$ for the normed space $(L^p(\mu)/\sim, \|\cdot\|_p)$.

We will abuse notation and continue to write members of $\mathcal{L}^p(\mu)$ as functions. Really they are equivalence classes, and so care must be taken when dealing with $\mathcal{L}^p(\mu)$. For example, if $f \in \mathcal{L}^p(\mu)$, it does *not* make sense to talk about the value of f at a point.

Theorem: Let (f_n) be a Cauchy sequence in $L^p(\mu)$. There exists $f \in L^p(\mu)$ with $\|f_n - f\|_p \to 0$. In fact, we can find a subsequence (n_k) such that $f_{n_k} \to f$ pointwise, almost everywhere.

[1] I have probably chosen the *opposite* notation from that used in most books, so be careful if you look in any textbook.

Proof: We shall show this carefully for $\mathcal{L}^\infty(\mu)$ below. This case is similar. $\square$

Corollary: $\mathcal{L}^p(\mu)$ is a Banach space.

Proposition: Let $(X, \mathcal{R}, \mu)$ be a measure space, and let $1 < p < \infty$. We can define a map $\Phi : \mathcal{L}^q(\mu) \to \mathcal{L}^p(\mu)^*$ by setting $\Phi(f) = F$, for $f \in \mathcal{L}^q(\mu)$, where

$$F : \mathcal{L}^p(\mu) \to \mathbb{K}, \quad g \mapsto \int_X fg \; d\mu \qquad (g \in \mathcal{L}^p(\mu)).$$

Proof: For $f \in L^p(\mu)$ and $g \in L^q(\mu)$, it follows by the above lemma (Holder's Inequality), that fg is integrable, and

$$\Big| \int_X fg \; d\mu \Big| \leq \int_X |fg| \; d\mu \leq \|f\|_q \|g\|_p.$$

Let $f_1, f_2 \in L^p(\mu)$ and $g_1, g_2 \in L^q(\mu)$ with $f_1 \sim f_2$ and $g_1 \sim g_2$. Then $f_1 g_1 = f_2 g_1$ almost everywhere and $f_2 g_1 = f_2 g_2$ almost everywhere, so $f_1 g_1 = f_2 g_2$ almost everywhere, and hence

$$\int_X f_1 g_1 \; d\mu = \int_X f_2 g_2 \; d\mu.$$

So Φ is well-defined.

Clearly Φ is linear, and we have shown that $\|\Phi(f)\| \leq \|f\|_q$.

Let $f \in L^q(\mu)$ and define $g : X \to \mathbb{K}$ by

$$g(x) = \begin{cases} \overline{f(x)} |f(x)|^{q-2} & : f(x) \neq 0, \\ 0 & : f(x) = 0. \end{cases}$$

Then $|g(x)| = |f(x)|^{q-1}$ for all $x \in X$, and so

$$\int_X |g|^p \; d\mu = \int_X |f|^{p(q-1)} \; d\mu = \int_X |f|^q \; d\mu,$$

so $\|g\|_p = \|f\|_q^{q/p}$, and so, in particular, $g \in \mathcal{L}^p(\mu)$. Let $F = \Phi(f)$, so that

$$F(g) = \int_X fg \; d\mu = \int_X |f|^q \; d\mu = \|f\|_q^q.$$

Thus $\|F\| \geq \|f\|_q^q / \|g\|_p = \|f\|_q$. So we conclude that $\|F\| = \|f\|_q$, showing that Φ is an isometry. $\square$

We will show that Φ is surjective, but this requires some more machinery.

Radon-Nikodym Theory

Let X be a set and let $\mathcal{R}$ be a σ-algebra on X. A *signed measure* on $\mathcal{R}$ is a map $\nu : \mathcal{R} \to \mathbb{R} \cup \{\pm\infty\}$ such that:

- $\nu(\emptyset) = 0$, and ν takes at most *one* of the values ∞ and $-\infty$;
- if $(A_n) \subseteq \mathcal{R}$ is a sequence of pairwise-disjoint sets, then $\nu\big(\bigcup_n A_n\big) = \sum_n \nu(A_n)$.

As ν takes only one of the values ∞ and $-\infty$, we never run into the problem of trying to evaluate $\infty - \infty$. The second condition is quite strong: it implies that the sum on the right must actually converge absolutely.

Theorem: Let ν be a signed measure. There exist $A, B \in \mathcal{R}$ with $A \cap B = \emptyset$, $A \cup B = X$ and such that for any $E \in \mathcal{R}$,

$$\nu(A \cap E) \geq 0, \quad \nu(B \cap E) \leq 0.$$

We call (A, B) a *Hahn-Decomposition* (of X) for ν. This need not be unique.

Proof: We only sketch this. We say that $A \in \mathcal{R}$ is *positive* if

$$\nu(E \cap A) \geq 0 \qquad (E \in \mathcal{R}),$$

and similiarly define what it means for a measurable set to be *negative*. Suppose that ν never takes the value $-\infty$ (the other case follows by considering the signed measure $-\nu$).

Let $\beta = \inf \nu(B_0)$ where we take the infimum over all negative sets B_0. If $\beta = -\infty$ then for each n, we can find a negative B_n with $\nu(B_n) \leq -n$. But then $B = \bigcup_n B_n$ would be negative with $\nu(B) \leq -n$ for any n, so that $\nu(B) = -\infty$ a contradiction.

So $\beta > -\infty$ and so for each n we can find a negative B_n $\nu(B_n) < \beta + 1/n$. Then we can show that $B = \bigcup_n B_n$ is negative, and argue that $\nu(B) \leq \beta$. As B is negative, actually $\nu(B) = \beta$.

There then follows a very tedious argument, by contradiction, to show that $A = X \setminus B$ is a positive set. Then (A, B) is the required decomposition. $\square$

Given a Hahn-Decomposition (A, B) for a signed measure ν, we define maps $\nu_+, \nu_- : \mathcal{R} \to [0, \infty]$ by

$$\nu_+(E) = \nu(A \cap E), \quad \nu_-(E) = -\nu(B \cap E) \qquad (E \in \mathcal{R}).$$

It is a simple check that ν_+ and ν_- are measures (and only one of them can take the value ∞). We can recover ν by observing that

$$\nu(E) = \nu\big((E \cap A) \cup (E \cap B)\big) = \nu(E \cap A) + \nu(E \cap B) = \nu_+(E) - \nu_-(E) \qquad (E \in \mathcal{R}).$$

So really signed measures are just the difference of two measures.

Let $|\nu| = \nu_+ + \nu_-$, so that $|\nu|$ is a measure. Given a measurable function $f : X \to \mathbb{K}$, we define

$$\int_X f \ d\nu = \int_X f \ d\nu_+ - \int_X f \ d\nu_- = \int_X \big(f\chi_A - f\chi_B\big) \ d|\nu|.$$

We can check that integration against ν behaves as we would expect (it is linear and so forth): this is most easily seen by looking the right hand side.

Let $(X, \mathcal{R}, \mu)$ be a measure space, and let ν be a signed measure defined on $\mathcal{R}$. We say that ν is *absolutely continuous* with respect to μ, written $\nu << \mu$, if whenever $\mu(E) = 0$, we have that $\nu(E) = 0$.

Proposition: Let $(X, \mathcal{R}, \mu)$ be a finite measure space, and let ν be a finite measure on $\mathcal{R}$, with $\nu << \mu$. Then there exists a measurable function $f : X \to [0, \infty)$ such that

$$\nu(E) = \int_E f \ d\mu = \int_X f\chi_E \ d\mu \qquad (E \in \mathcal{R}).$$

Proof: Again, this is a sketch. Let $\mathcal{D}$ be the collection of measurable functions $g : X \to [0, \infty)$ such that

$$\int_E g \ d\mu \leq \nu(E) \qquad (E \in \mathcal{R}).$$

Let $\alpha = \sup_{g \in \mathcal{D}} \int_X g \ d\mu \leq \nu(X) < \infty$. So we can find a sequence (g_n) in $\mathcal{D}$ with $\int_X g_n \ d\mu \to \alpha$.

We define $f_0(x) = \sup_n g_n(x)$. We can show that $f_0 = \infty$ only on a set of μ-measure zero, so if we adjust f_0 on this set, we get a measurable function $f : X \to [0, \infty)$. There is now a long argument to show that f is as required. $\square$

Theorem (Radon-Nikodym Theorem): Let $(X, \mathcal{R}, \mu)$ be a finite measure space, let ν be a signed measure on $\mathcal{R}$ with $\nu << \mu$. Let $\nu = \nu_+ - \nu_-$ be the decomposition given a Hahn-Decomposition for ν. If $|\nu|$ is a finite measure, then there exists a measurable function $f : X \to \mathbb{R}$ such that

$$\nu(E) = \int_E f \; d\mu \qquad (E \in \mathcal{R}).$$

If g is another function with this property, then $f = g$ almost everywhere.

Proof: We can find f by applying the previous operation to the measures ν_+ and ν_- (as it is easy to verify that $\nu_+, \nu_- << \mu$).

We show that f is essentially unique. If g is another function inducing ν, then

$$\int_E f - g \; d\mu = \nu(E) - \nu(E) = 0 \qquad (E \in \mathcal{R}).$$

Let $E = \{x \in X : f(x) - g(x) \geq 0\}$, so as $f - g$ is measurable, $E \in \mathcal{R}$. Then $\int_E f - g \; d\mu = 0$ and $f - g \geq 0$ on E, so by our result from integration theory, we have that $f - g = 0$ almost everywhere on E. Similarly, if $F = \{x \in X : f(x) - g(x) \leq 0\}$, then $F \in \mathcal{R}$ and $f - g = 0$ almost everywhere on F. As $E \cup F = X$, we conclude that $f = g$ almost everywhere. $\square$

We now briefly discuss complex measures. Let X be a set and $\mathcal{R}$ be a σ-algebra on X. A *complex measure* is a map $\mu : \mathcal{R} \to \mathbb{C}$ such that, if we define

$$\mu_r(E) = \Re\mu(E), \quad \mu_i(E) = \Im\mu(E) \qquad (E \in \mathcal{R}),$$

then μ_r and μ_i are signed measures.[2] When $f : X \to \mathbb{C}$ is measurable, and integrable for μ_r and μ_i, we define

$$\int_X f \; d\mu = \int_X f \; d\mu_r + i \int_X f \; d\mu_i.$$

Theorem (Complex Radon-Nikodym Theorem): Let $(X, \mathcal{R}, \mu)$ be a finite measure space, let ν be a complex measure on $\mathcal{R}$ with $\nu << \mu$. Suppose that $|\nu_r|$ and $|\nu_i|$ are finite measures.[3] There exists a measurable function $f : X \to \mathbb{C}$ such that

$$\nu(E) = \int_E f \; d\mu \qquad (E \in \mathcal{R}).$$

Again, if g is another function with this property, then $f = g$ almost everywhere.

Proof: Just take real and imaginary parts, and apply the main Radon-Nikodym Theorem. $\square$

Application to $\mathcal{L}^p$ spaces

Proposition: Let $(X, \mathcal{R}, \mu)$ be a finite measure space, let $1 < p < \infty$, and let $F \in \mathcal{L}^p(\mu)^*$. Then there exists $g \in \mathcal{L}^q(\mu)$ such that

$$F(f) = \int_X fg \; d\mu \qquad (f \in \mathcal{L}^p(\mu)).$$

[2]Notice that this definition does not allow a notion of "infinity", as μ takes values in $\mathbb{C}$. It would be possible to somehow handle infinity, but for us, it will never be a problem.

[3]Which is actually automatic for our definition of what a complex measure is!

Proof: As $\mu(X) < \infty$, for $E \in \mathcal{R}$, we have that $\|\chi_E\|_p = \mu(E)^{1/p} < \infty$. So $\chi_E \in \mathcal{L}^p(\mu)$, and hence we can define

$$\nu(E) = F(\chi_E) \qquad (E \in \mathcal{R}).$$

We proceed to show that ν is a signed (or complex) measure. Then we can apply the Radon-Nikodym Theorem to find a function $g : X \to \mathbb{K}$ such that

$$F(\chi_E) = \nu(E) = \int_E g \, d\mu \qquad (E \in \mathcal{R}).$$

There is then a long argument to show that $g \in L^q(\mu)$ and that

$$\int_X fg \, d\mu = F(f)$$

for all $f \in \mathcal{L}^p(\mu)$, and not just for $f = \chi_E$. $\square$

Corollary: For $1 < p < \infty$, we have that $\mathcal{L}^p(\mu)^* = \mathcal{L}^q(\mu)$ isometrically, under the identification of the above results.

Proposition: Let $(X, \mathcal{R}, \mu)$ be a finite measure space, and let $1 \leq p < \infty$. Then the collection of simple functions is dense in $\mathcal{L}^p(\mu)$.

Proof:[4] Let $f \in \mathcal{L}^p(\mu)$, and suppose for now that $f \geq 0$. For each $n \in \mathbb{N}$, let

$$f_n = \max(n, 2^{-n}\lfloor 2^n f \rfloor).$$

Then each f_n is simple, $f_n \uparrow f$, and $|f_n - f|^p \to 0$ pointwise. For each n, we have that

$$0 \leq f_n \leq f \implies 0 \leq f - f_n \leq f,$$

so that $|f - f_n|^p \leq |f|^p$ for all n. As $\int |f|^p \, d\mu < \infty$, we can apply the Dominated Convergence Theorem to see that

$$\lim_n \int_X |f_n - f|^p \, d\mu = 0,$$

that is, $\|f_n - f\|_p \to 0$.

The general case follows by taking positive and negative parts, and if $\mathbb{K} = \mathbb{C}$, by taking real and imaginary parts first. $\square$

Proposition: Let $([0,1], \mathcal{R}, \mu)$ be the restriction of Lebesgue measure to $[0,1]$. We often write $\mathcal{L}^p([0,1])$ instead of $\mathcal{L}^p(\mu)$. For $1 < p < \infty$, we have that $C_\mathbb{K}([0,1])$ is dense in $\mathcal{L}^p([0,1])$.

Proof: As $[0,1]$ is a finite measure space, and each member of $C_\mathbb{K}([0,1])$ is bounded, it is easy to see that each $f \in C_\mathbb{K}([0,1])$ is such that $\|f\|_p < \infty$. So it makes sense to regard $C_\mathbb{K}([0,1])$ as a subspace of $\mathcal{L}^p(\mu)$. If $C_\mathbb{K}([0,1])$ is not dense in $\mathcal{L}^p(\mu)$, then we can find a non-zero $F \in \mathcal{L}^p([0,1])^*$ with $F(f) = 0$ for each $f \in C_\mathbb{K}([0,1])$. This was a corollary of the Hahn-Banach theorem which we proved in Chapter 1.

So there exists a non-zero $g \in \mathcal{L}^q([0,1])$ with

$$\int_{[0,1]} fg \, d\mu = 0 \qquad (f \in C_\mathbb{K}([0,1])).$$

Let $a < b$ in $[0,1]$. By approximating $\chi_{(a,b)}$ by a continuous function, we can show that $\int_{(a,b)} g \, d\mu = \int g\chi_{(a,b)} \, d\mu = 0$.

[4]In lectures, we actually needed this argument in proving the above theorem that $\mathcal{L}^p(\mu)^* = \mathcal{L}^q(\mu)$.

Suppose for now that $\mathbb{K} = \mathbb{R}$. Let $A = \{x \in [0,1] : g(x) \geq 0\} \in \mathcal{R}$. By the definition of the Lebesgue (outer) measure, for $\epsilon > 0$, there exist sequences (a_n) and (b_n) with $A \subseteq \bigcup_n (a_n, b_n)$, and $\sum_n (b_n - a_n) \leq \mu(A) + \epsilon$.

For each N, consider $\bigcup_{n=1}^N (a_n, b_n)$. If some (a_i, b_i) overlaps (a_j, b_j), then we could just consider the larger interval $(\min(a_i, a_j), \max(b_i, b_j))$. Formally by an induction argument, we see that we can write $\bigcup_{n=1}^N (a_n, b_n)$ as a finite union of *disjoint* open intervals. By linearity, it hence follows that for $N \in \mathbb{N}$, if we set $B_N = \bigcup_{n=1}^N (a_n, b_n)$, then

$$\int g\chi_{B_N} \, d\mu = \int g\chi_{(a_1,b_1)\cup\cdots\cup(a_N,b_N)} \, d\mu = 0.$$

Let $B = \bigcup_n (a_n, b_n)$, so $A \subseteq B$ and $\mu(B) \leq \sum_n (b_n - a_n) \leq \mu(A) + \epsilon$. We then have that

$$\Big| \int g\chi_{B_N} \, d\mu - \int g\chi_B \, d\mu \Big| = \Big| \int g\chi_{B\setminus(a_1,b_1)\cup\cdots\cup(a_N,b_N)} \, d\mu \Big|.$$

We now apply the Holder inequality to get

$$\leq \Big(\int \chi_{B\setminus(a_1,b_1)\cup\cdots\cup(a_N,b_N)} \, d\mu \Big)^{1/p} \|g\|_q = \mu(B \setminus (a_1,b_1) \cup \cdots \cup (a_N,b_N))^{1/p} \|g\|_q$$
$$\leq \Big(\sum_{n=N+1}^{\infty} (b_n - a_n) \Big)^{1/p} \|g\|_q.$$

We can make this arbitrarily small by making N large. Hence we conclude that

$$\int g\chi_B \, d\mu = 0.$$

Then we apply Holder again to see that

$$\Big| \int g\chi_A \, d\mu \Big| = \Big| \int g\chi_A \, d\mu - \int g\chi_B \, d\mu \Big| = \Big| \int g\chi_{B\setminus A} \, d\mu \Big| \leq \|g\|_q \mu(B \setminus A)^{1/p} \leq \|g\|_q \epsilon^{1/p}.$$

As $\epsilon > 0$ was arbitrary, we see that $\int_A g \, d\mu = 0$. As g is positive on A, we conclude that $g = 0$ almost everywhere on A.

A similar argument applied to the set $\{x \in [0,1] : g(x) \leq 0\}$ allows us to conclude that $g = 0$ almost everywhere. If $\mathbb{K} = \mathbb{C}$, then take real and imaginary parts. $\square$

We now turn our attention to $\mathcal{L}^1(\mu)$, and its dual space. As a *warning*, for some of what follows, it is necessary (and not just a simplification) to consider *finite* measure spaces.[5] You may also find different definitions in books, although for finite measure spaces, these will boil down to being the same as our definitions.

Let $(X, \mathcal{R}, \mu)$ be a measure space. A measurable function $f : X \to \mathbb{K}$ is *essentially bounded* if there exists $K > 0$ such that $|f| \leq K$ almost everywhere. We set

$$\text{ess-sup}_X |f| = \inf \big\{ K > 0 : |f| \leq K \text{ almost everywhere} \big\}.$$

Lemma: For an essentially bounded $f : X \to \mathbb{K}$, let $K = \text{ess-sup}_X |f|$. Then $|f| \leq K$ almost everywhere.

Proof: By definition, if we set

$$A_n = \{x \in X : |f(x)| > K + 1/n\} \qquad (n \in \mathbb{N}),$$

[5]Technically, everything will work for even σ-finite measures, but things do generally go wrong for arbitrary measures.

then $\mu(A_n) = 0$. Hence, if $A = \bigcup_n A_n$, then $\mu(A) = 0$. If $|f(x)| > K$ then for some n, we have that $|f(x)| > K + 1/n$, so that $x \in A$. It follows that $A = \{x \in X : |f(x)| > K\}$, and so $|f| \leq K$ almost everywhere. $\square$

We let $L^\infty(\mu)$ be the collection of all essentially bounded functions $f : X \to \mathbb{K}$. It is easy to see that this is a vector space. We define $\|\cdot\|_\infty$ on $L^\infty(\mu)$ by setting

$$\|f\|_\infty = \text{ess-sup}_X |f| \qquad (f \in \ell^\infty(\mu)).$$

Proposition: $\|\cdot\|_\infty$ is a norm on $L^\infty(\mu)/\sim$.

Proof: If $f = g$ almost everywhere and $|f| \leq K$ almost everywhere, then $|g| \leq K$ almost everywhere. So it follows that $\|f\|_\infty = \|g\|_\infty$. Hence $\|\cdot\|_\infty$ is well-defined on $\ell^\infty(\mu)/\sim$. Notice that for $f \in L^\infty(\mu)$, $\|f\|_\infty = 0$ if and only if $|f| = 0$ almost everywhere. So $L^\infty(\mu)/\sim$ is a vector space.

If $f, g \in L^\infty(\mu)$ and $a \in \mathbb{K}$, then clearly $\|af\|_\infty = |a|\|f\|_\infty$. If $|f| \leq K$ almost everywhere, and $|g| \leq L$ almost everywhere, then it is easy to see that $|f + g| \leq K + L$. It follows that $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. $\square$

As before, we write $\mathcal{L}^\infty(\mu)$ for $L^\infty(\mu)/\sim$.

Theorem: For a measure space $(X, \mathcal{R}, \mu)$, we have that $\mathcal{L}^\infty(\mu)$ is a Banach space.

Proof: Again, it is enough to work in $L^\infty(\mu)$. Let (f_n) be a Cauchy sequence in $L^\infty(\mu)$. By moving to a subsequence if necessary, we may suppose that $\|f_n - f_{n+1}\|_\infty \leq 2^{-n}$ for each n. For each n, let

$$A_n = \{x \in X : |f_n - f_{n+1}| > 2^{-n}\} \in \mathcal{R}.$$

By the definition of $\|\cdot\|_\infty$, we see that each A_n is a null set. Hence $A = \bigcup_n A_n$ is also null. Define

$$\tilde{f}_n(x) = \begin{cases} f_n(x) & : x \notin A, \\ 0 & : x \in A, \end{cases} \qquad (n \in \mathbb{N}).$$

Hence $f_n \sim \tilde{f}_n$ for each n. For $N > 0$ and $x \in X$, we see that

$$\sum_{n \geq N} |\tilde{f}_n(x) - \tilde{f}_{n+1}(x)| \leq \sum_{n \geq N} 2^{-n} = 2^{1-N}.$$

So the sum $\sum_n \tilde{f}_n(x) - \tilde{f}_{n+1}(x)$ converges absolutely, and so converges in $\mathbb{K}$, say to $g(x)$. However, notice that

$$\tilde{f}_1(x) - \Big(\sum_{n=1}^{N-1} \tilde{f}_n(x) - \tilde{f}_{n+1}(x) \Big) = \tilde{f}_N(x).$$

So we see that $\tilde{f}_N(x) \to \tilde{f}_1(x) - g(x)$ for each $x \in X$. Define $f(x) = \tilde{f}_1(x) - g(x)$ for each $x \in X$.

We skip showing that g (and hence f) is measurable (we have seen this proof before, and you could just state it without proof in an exam).

For $x \in X$, we have that

$$\begin{aligned} |f(x)| = |\tilde{f}_1(x) - g(x)| &= \lim_N |\tilde{f}_N(x)| = \lim_N \Big| \tilde{f}_1(x) - \Big(\sum_{n=1}^{N-1} \tilde{f}_n(x) - \tilde{f}_{n+1}(x) \Big) \Big| \\ &\leq |\tilde{f}_1(x)| + \sum_{n=1}^{\infty} 2^{-n} = 1 + |\tilde{f}_1(x)|. \end{aligned}$$

Hence $|f| \leq 1 + |f_1|$ almost everywhere (as $\tilde{f}_1 \sim f_1$) and so $f \in L^\infty(\mu)$.

Finally, for $x \in X$, we see that for $r \geq 1$,

$$|f(x) - \tilde{f}_r(x)| = \lim_N |\tilde{f}_N(x) - \tilde{f}_r(x)| = \lim_N \Big| \sum_{n=r}^{N-1} \tilde{f}_n(x) - \tilde{f}_{n+1}(x) \Big| \leq \lim_N \sum_{n=r}^{N-1} 2^{-n} = 2^{1-r}.$$

So we see that $\|f - \tilde{f}_r\|_\infty \to 0$, and as $f_r \sim \tilde{f}_r$ for each r, we also have that $\|f - f_r\|_\infty \to 0$, as required. $\square$

Notice that we haven't yet really used that our measure space is finite. The next proof changes this.

Proposition: Let $(X, \mathcal{R}, \mu)$ be a finite measure space. We can define a map $\Phi : \mathcal{L}^\infty(\mu) \to \mathcal{L}^1(\mu)^*$ by setting, for $f \in \mathcal{L}^\infty(\mu)$, $\Phi(f) = F$, where

$$F(g) = \int_X fg \ d\mu \qquad (g \in \mathcal{L}^1(\mu)).$$

Then Φ is a linear map, which is an isometry.

Proof: As in the analogous proof for $\mathcal{L}^p$, we can check that Φ is well-defined on equivalence classes. We note that if $|f| \leq K$ almost everywhere, then

$$\Big| \int_X fg \ d\mu \Big| \leq \int_X |fg| \ d\mu \leq K \int_X |g| \ d\mu = K\|g\|_1.$$

So the integral is defined, and it hence follows that $|F(g)| \leq \|g\|_1 \|f\|_\infty$. Clearly Φ is linear, and we just showed that $\|\Phi\| \leq 1$.

Fix $f \in \mathcal{L}^\infty(\mu)$. If $\|f\|_\infty = 0$ then clearly $F = 0$. Otherwise, let $\epsilon > 0$ be such that $\|f\|_\infty - \epsilon > 0$. Then we see that

$$A = \{x \in X : |f(x)| \geq \|f\|_\infty - \epsilon\}$$

is not null. Define $g : X \to \mathbb{K}$ by

$$g(x) = \begin{cases} \overline{f(x)}|f(x)|^{-1} & : x \in A, \\ 0 & : x \notin A. \end{cases}$$

We have chosen this g because $g(x)f(x) = \chi_A(x)|f(x)|$ for all $x \in X$. Notice first that

$$\int_X |g| \ d\mu = \int_A 1 \ d\mu = \mu(A) < \infty,$$

so $g \in \mathcal{L}^1(\mu)$ with $\|g\|_1 = \mu(A)$. Then

$$|F(g)| = \Big| \int_X fg \ d\mu \Big| = \int_A |f| \ d\mu \geq \mu(A)\big(\|f\|_\infty - \epsilon\big) = \|g\|_1\big(\|f\|_\infty - \epsilon\big).$$

Hence $\|F\| \geq \|f\|_\infty - \epsilon$. As $\epsilon > 0$ was arbitrary, we conclude that $\|F\| = \|f\|_\infty$, so Φ is an isometry. $\square$

Theorem: With the notation of the previous proposition, Φ is surjective. That is, for each $F \in \mathcal{L}^1(\mu)^*$, there exists $f \in \mathcal{L}^\infty(\mu)$ with

$$F(g) = \int_X fg \ d\mu \qquad (g \in \mathcal{L}^1(\mu)).$$

Proof: We define $\lambda : \mathcal{R} \to \mathbb{K}$ by

$$\lambda(A) = F(\chi_A).$$

This makes sense, as $\|\chi_A\|_1 = \mu(A) < \infty$. Then $\lambda(\emptyset) = 0$. For (A_n) a sequence of pairwise disjoint sets in $\mathcal{R}$, let $A = \bigcup_n A_n$, so that

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(A) \leq \mu(X) < \infty.$$

Hence the sum converges. We also see that

$$\begin{aligned}\left\|\chi_A - \chi_{A_1\cup\cdots\cup A_n}\right\|_1 &= \left\|\chi_{A_{n+1}\cup A_{n+2}\cup\cdots}\right\|_1 \\ &= \int \chi_{A_{n+1}\cup A_{n+2}\cup\cdots} \ d\mu = \mu(A_{n+1} \cup A_{n+2} \cup \cdots) = \sum_{k>n} \mu(A_k),\end{aligned}$$

which tends to 0 as n tends to infinity. So as F is continuous,

$$\begin{aligned}\lambda(A) = F(\chi_A) &= \lim_n F(\chi_{A_1\cup\cdots\cup A_n}) = \lim_n F(\chi_{A_1} + \cdots + \chi_{A_n}) \\ &= \lim_n \sum_{k=1}^{n} F(\chi_{A_k}) = \sum_{k=1}^{\infty} \lambda(A_k).\end{aligned}$$

So λ is a signed or complex measure.

Clearly, if $\mu(A) = 0$, then $\chi_A = 0$ in $\mathcal{L}^1(\mu)$, and so $\lambda(A) = 0$. Hence $\lambda << \mu$, so by Radon-Nikodym, there exists $f : X \to \mathbb{K}$ measurable with

$$F(\chi_A) = \lambda(A) = \int_A f \ d\mu = \int_X f\chi_A \ d\mu \qquad (A \in \mathcal{R}).$$

Suppose that $\mathbb{K} = \mathbb{R}$. For $K > 0$ let $A = \{x \in X : f(x) > K\} \in \mathcal{R}$. Then

$$K\mu(A) < \int_A f \ d\mu = F(\chi_A) \leq \|F\|\|\chi_A\|_1 = \|F\|\mu(A).$$

If A is not null, then $K < \|F\|$. So if $K = \|F\|$, then A must be null, and so we conclude that $f \leq \|F\|$ almost everywhere. We can similarly show that $f \geq -\|F\|$ almost everywhere. So $f \in \mathcal{L}^\infty(\mu)$. If $\mathbb{K} = \mathbb{C}$, then we take real and imaginary parts (but note that this won't give a perfect estimate of $\|f\|_\infty$, but it will show that $f \in \mathcal{L}^\infty(\mu)$).

As the linear span of indicator functions is the collection of simple functions, we immediately see that if $g \in \mathcal{L}^1(\mu)$ is a simple function, then

$$F(g) = \int_X fg \ d\mu = \Phi(f)(g).$$

As we showed above that simple functions are dense in $\mathcal{L}^1(\mu)$, and both F and $\Phi(f)$ are continuous, we conclude that $F = \Phi(f)$ on all of $\mathcal{L}^1(\mu)$. As Φ is an isometry, we also have that $\|f\|_\infty = \|F\|$. $\square$

We now apply this result.

Proposition: We have that $C_{\mathbb{K}}([0,1])$ is dense in $\mathcal{L}^1([0,1])$.

Proof: This follows exactly as in the $\mathcal{L}^p([0,1])$ case, now that we know that $\mathcal{L}^1([0,1])^* = \mathcal{L}^\infty([0,1])$. $\square$

Finally, we recall that on the example sheet, we showed that $C_{\mathbb{K}}([0,1])$ is *not* dense in $\mathcal{L}^\infty([0,1])$.

Linear Analysis IV: The dual of $C(K)$

Updated: 3rd December 2008

Let K be a compact (always assumed Hausdorff) topological space. The *Borel σ-algebra*, $\mathcal{B}(K)$, on K, is the σ-algebra generated by the open sets in K (recall what this means from Chapter 2). A member of $\mathcal{B}(K)$ is a *Borel* set.

Notice that if $f : K \to \mathbb{K}$ is a continuous function, then clearly f is $\mathcal{B}(K)$-measurable (the inverse image of an open set will be open, and hence certainly Borel). So if $\mu : \mathcal{B}(K) \to \mathbb{K}$ is a finite signed or complex measure (for $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ respectively), then f will be μ-integrable (as f is bounded) and so we can define

$$\phi_\mu : C_{\mathbb{K}}(K) \to \mathbb{K}, \quad \phi_\mu(f) = \int_K f \; d\mu \qquad (f \in C_{\mathbb{K}}(K)).$$

Clearly ϕ_μ is linear. Suppose for now that μ is positive, so that

$$|\phi_\mu(f)| \leq \int_K |f| \; d\mu \leq \|f\|_\infty \mu(K) \qquad (f \in C_{\mathbb{K}}(K)).$$

So $\phi_\mu \in C_{\mathbb{K}}(K)^*$ with $\|\phi_\mu\| \leq \mu(K)$.

The aim of this chapter is to show that all of $C_{\mathbb{K}}(K)^*$ arises in this way.

A measure $\mu : \mathcal{B}(K) \to [0, \infty)$ is *regular* if for each $A \in \mathcal{B}(K)$, we have

$$\begin{aligned} \mu(A) &= \sup \big\{ \mu(E) : E \subseteq A \text{ and } E \text{ is compact} \big\} \\ &= \inf \big\{ \mu(U) : A \subseteq U \text{ and } U \text{ is open} \big\}. \end{aligned}$$

As we are working only with compact spaces, for us, "compact" is the same as "closed". A signed measure ν is *regular* if ν_+ and ν_- are signed measures. A complex measure is *regular* if its real and imaginary parts are regular.

Regular measures somehow interact "well" with the underlying topology on K.

We let $M_{\mathbb{R}}(K)$ and $M_{\mathbb{C}}(K)$ be the collection of all finite, regular, signed or complex (respectively) measures on $\mathcal{B}(K)$. These are real or complex, respectively, vector spaces for the obvious definition of addition and scalar multiplication.

For $\mu \in M_{\mathbb{K}}(K)$ we define

$$\|\mu\| = \sup \Big\{ \sum_{n=1}^\infty |\mu(A_n)| \Big\},$$

where the supremum is taken over all sequences (A_n) of pairwise disjoint members of $\mathcal{B}(K)$, with $\bigcup_n A_n = K$. Such (A_n) are called *partitions*.

Proposition: $\|\cdot\|$ is a norm on $M_{\mathbb{K}}(K)$.

Proof: If $\mu = 0$ then clearly $\|\mu\| = 0$. If $\|\mu\| = 0$, then for $A \in \mathcal{B}(K)$, let $A_1 = A, A_2 = K \setminus A$ and $A_3 = A_4 = \cdots = \emptyset$. Then (A_n) is a partition, and so

$$0 = \sum_{n=1}^\infty |\mu(A_n)| = |\mu(A)| + |\mu(K \setminus A)|.$$

Hence $\mu(A) = 0$, and so as A was arbitrary, we have that $\mu = 0$.

Clearly $\|a\mu\| = |a|\|\mu\|$ for $a \in \mathbb{K}$ and $\mu \in M_{\mathbb{K}}(K)$.

For $\mu, \lambda \in M_{\mathbb{K}}(K)$ and a partition (A_n), we have that

$$\sum_n |(\mu + \lambda)(A_n)| = \sum_n |\mu(A_n) + \lambda(A_n)| \leq \sum_n |\mu(A_n)| + |\lambda(A_n)| \leq \|\mu\| + \|\lambda\|.$$

As (A_n) was arbitrary, we see that $\|\mu + \lambda\| \leq \|\mu\| + \|\lambda\|$. $\square$

To get a handle on the "regular" condition, we need to know a little more about $C_{\mathbb{K}}(K)$.

Theorem (Urysohn's Lemma): Let K be a compact space, and let E, F be closed subsets of K with $E \cap F = \emptyset$. There exists $f : K \to [0,1]$ continuous with $f(x) = 1$ for $x \in E$ and $f(x) = 0$ for $x \in F$ (written $f(E) = \{1\}$ and $f(F) = \{0\}$).

Proof: See a book on (point set) topology. $\square$

Lemma: Let $\mu : \mathcal{B}(K) \to [0, \infty)$ be a regular measure. Then for $U \subseteq K$ open, we have

$$\mu(U) = \sup\Big\{ \int_K f \ d\mu : f \in C_{\mathbb{R}}(K), 0 \leq f \leq \chi_U \Big\}.$$

Proof: If $0 \leq f \leq \chi_U$, then

$$0 = \int_K 0 \ d\mu \leq \int_K f \ d\mu \leq \int_K \chi_U \ d\mu = \mu(U).$$

Conversely, let $F = K \setminus U$, a closed set. Let $E \subseteq U$ be closed. By Urysohn, there exists $f : K \to [0,1]$ continuous with $f(E) = \{1\}$ and $f(F) = \{0\}$. So $\chi_E \leq f \leq \chi_U$, and hence

$$\mu(E) \leq \int_K f \ d\mu \leq \mu(U).$$

As μ is regular,

$$\mu(U) = \sup\big\{\mu(E) : E \subseteq U \text{ closed}\big\} \leq \sup\Big\{ \int_K f \ d\mu : 0 \leq f \leq \chi_U \Big\} \leq \mu(U).$$

Hence we have equality throughout. $\square$

Lemma: Let $\mu \in M_{\mathbb{R}}(K)$. Then

$$\|\mu\| = \|\phi_\mu\| := \sup\Big\{ \Big| \int_K f \ d\mu \Big| : f \in C_{\mathbb{R}}(K), \|f\|_\infty \leq 1 \Big\}.$$

Proof: Let (A, B) be a Hahn-Decomposition for μ. For $f \in C_{\mathbb{R}}(K)$ with $\|f\|_\infty \leq 1$, we have that

$$\begin{aligned}\Big| \int_K f \ d\mu \Big| &\leq \Big| \int_A f \ d\mu \Big| + \Big| \int_B f \ d\mu \Big| = \Big| \int_A f \ d\mu_+ \Big| + \Big| \int_B f \ d\mu_- \Big| \\ &\leq \int_A |f| \ d\mu_+ + \int_B |f| \ d\mu_- \leq \|f\|_\infty \big(\mu(A) - \mu(B)\big) \leq \|f\|_\infty \|\mu\|,\end{aligned}$$

using the fact that $\mu(B) \leq 0$ and that (A, B) is a partition of K.

Conversely, as μ is regular, for $\epsilon > 0$, there exist closed sets E and F with $E \subseteq A$, $F \subseteq B$, and with $\mu_+(E) > \mu_+(A) - \epsilon$ and $\mu_-(F) > \mu_-(B) - \epsilon$. By Urysohn, there exists $f : K \to [0,1]$ continuous with $f(E) = \{1\}$ and $f(F) = \{0\}$. Let $g = 2f - 1$, so g is continuous, g takes values in $[-1, 1]$, and $g(E) = \{1\}$, $g(F) = \{-1\}$. Then

$$\begin{aligned}\int_K g \ d\mu &= \int_E 1 \ d\mu + \int_F -1 \ d\mu + \int_{K \setminus (E \cup F)} g \ d\mu \\ &= \mu(E) - \mu(F) + \int_{A \setminus E} g \ d\mu + \int_{B \setminus F} g \ d\mu\end{aligned}$$

As $E \subseteq A$, we have $\mu(E) = \mu_+(E)$, and as $F \subseteq B$, we have $-\mu(F) = \mu_-(F)$. So

$$\begin{aligned}\int_K g\ d\mu &> \mu_+(A) - \epsilon + \mu_-(B) - \epsilon + \int_{A\setminus E} g\ d\mu + \int_{B\setminus F} g\ d\mu \\ &\geq |\mu(A)| + |\mu(B)| - 2\epsilon - |\mu(A \setminus E)| - |\mu(B \setminus F)| \\ &\geq |\mu(A)| + |\mu(B)| - 4\epsilon.\end{aligned}$$

As $\epsilon > 0$ was arbitrary, we see that $\|\phi_\mu\| \geq |\mu(A)| + |\mu(B)|$.

Finally, let (A_n) be a partition of K. Then

$$\begin{aligned}\sum_n |\mu(A_n)| &= \sum_n |\mu(A_n \cap A) + \mu(A_n \cap B)| \leq \sum_n |\mu(A_n \cap A)| + |\mu(A_n \cap B)| \\ &= \sum_n \mu(A_n \cap A) - \mu(A_n \cap B) = \mu(A) - \mu(B) = |\mu(A)| + |\mu(B)|.\end{aligned}$$

So $\|\mu\| = |\mu(A)| + |\mu(B)|$, finishing the proof. □

We shall deal with the complex case later.

The following is the key point of this chapter.

Theorem (Riesz Representation): Let K be a compact (Hausdorff) space, and let $\lambda \in C_{\mathbb{K}}(K)^*$. There exists a unique $\mu \in M_{\mathbb{K}}(K)$ such that

$$\lambda(f) = \int_K f\ d\mu \qquad (f \in C_{\mathbb{K}}(K)).$$

Furthermore, $\|\lambda\| = \|\mu\|$.

Proof: Let us show uniqueness. If $\mu_1, \mu_2 \in M_{\mathbb{K}}(K)$ both induce λ then $\mu = \mu_1 - \mu_2$ induces the zero functional on $C_{\mathbb{K}}(K)$. So for $f \in C_{\mathbb{R}}(K)$,

$$\begin{aligned}0 = \Re \int_K f\ d\mu &= \int_K f\ d\mu_r \\ = \Im \int_K f\ d\mu &= \int_K f\ d\mu_i.\end{aligned}$$

So μ_r and μ_i both induce the zero functional on $C_{\mathbb{R}}(K)$. By the previous lemma, this means that $\|\mu_r\| = \|\mu_i\| = 0$, showing that $\mu = \mu_r + i\mu_i = 0$, as required.

Existence is harder, and we shall only sketch it here. Firstly, we shall suppose that $\mathbb{K} = \mathbb{R}$ and that λ is *positive*, that is, $\lambda(f) \geq 0$ if $f \in C_{\mathbb{R}}(K)$ with $f \geq 0$. We now need a technical definition. For $f \in C_{\mathbb{R}}(K)$, we define the *support* of f, written $\operatorname{supp}(f)$, to be the closure of the set $\{x \in K : f(x) \neq 0\}$.

Motivated by the above lemmas, for $U \subseteq K$ open, we define

$$\mu^*(U) = \sup\Big\{\lambda(f) : f \in C_{\mathbb{R}}(K), 0 \leq f \leq \chi_U, \operatorname{supp}(f) \subseteq U\Big\}.$$

For $A \subseteq K$ general, we define

$$\mu^*(A) = \inf\big\{\mu^*(U) : U \subseteq K \text{ is open}, A \subseteq U\big\}.$$

We then proceed to show that μ^* is an outer measure: this requires a technical topological lemma, where we make use of the support condition in the definition. We then check that every open set in μ^*-measurable. As $\mathcal{B}(K)$ is generated by open sets, and the collection of μ^*-measurable sets is a σ-algebra, it follows that every member of $\mathcal{B}(K)$ is μ^*-measurable. By using results from Chapter 2, it follows that if we let μ be the restriction of μ^* to $\mathcal{B}(K)$, then μ is a measure on $\mathcal{B}(K)$. We then check that this measure is regular. Finally, we show that μ does induce the functional λ. Arguably, it is this last step which is the hardest (or least natural to prove).

If λ is not positive, then for $f \in C_{\mathbb{R}}(K)$ with $f \geq 0$, we define

$$\begin{aligned}
\lambda_+(f) &= \sup\big\{\lambda(g) : g \in C_{\mathbb{R}}(K), 0 \leq g \leq f\big\} \geq 0, \\
\lambda_-(f) &= \lambda_+(f) - \lambda(f) = \sup\big\{\lambda(g) - \lambda(f) : g \in C_{\mathbb{R}}(K), 0 \leq g \leq f\big\} \\
&= \sup\big\{\lambda(h) : h \in C_{\mathbb{R}}(K), 0 \leq h + f \leq f\big\} \\
&= \sup\big\{\lambda(h) : h \in C_{\mathbb{R}}(K), -f \leq h \leq 0\big\} \geq 0.
\end{aligned}$$

We can check that

$$\lambda_+(tf) = t\lambda_+(f), \quad \lambda_-(tf) = t\lambda_-(f) \qquad (t \geq 0, f \geq 0).$$

For $f_1, f_2 \geq 0$, we have that

$$\begin{aligned}
\lambda_+(f_1 + f_2) &= \sup\big\{\lambda(g) : 0 \leq g \leq f_1 + f_2\big\} \\
&= \sup\big\{\lambda(g_1 + g_2) : 0 \leq g_1 + g_2 \leq f_1 + f_2\big\} \\
&\geq \sup\big\{\lambda(g_1) + \lambda(g_2) : 0 \leq g_1 \leq f_1, 0 \leq g_2 \leq f_1 + f_2\big\} = \lambda_+(f_1) + \lambda_+(f_2).
\end{aligned}$$

Conversely, if $0 \leq g \leq f_1 + f_2$, then set $g_1 = \min(g, f_1)$, so $0 \leq g_1 \leq f_1$. Let $g_2 = g - g_1$ so $g_1 \leq g$ implies that $0 \leq g_2$. For $x \in K$, if $g_1(x) = g(x)$ then $g_2(x) = 0 \leq f_2(x)$; if $g_1(x) = f_1(x)$ then $f_1(x) \leq g(x)$ and so $g_2(x) = g(x) - f_1(x) \leq f_2(x)$. So $0 \leq g_2 \leq f_2$, and $g = g_1 + g_2$. So in the above displayed equation, we really have equality throughout, and so $\lambda_+(f_1 + f_2) = \lambda_+(f_1) + \lambda_+(f_2)$. As λ is additive, it is now immediate that $\lambda_-(f_1 + f_2) = \lambda_-(f_1) + \lambda_-(f_2)$

For $f \in C_{\mathbb{R}}(K)$ we define

$$\lambda_+(f) = \lambda_+(f_+) - \lambda_+(f_-), \quad \lambda_-(f) = \lambda_-(f_+) - \lambda_-(f_-).$$

As when we were dealing with integration, we can check that λ_+ and λ_- become linear functionals; it is easy to see that they are bounded. As λ_+ and λ_- are positive functionals, we can find μ_+ and μ_- positive measures in $M_{\mathbb{R}}(K)$ such that

$$\lambda_+(f) = \int_K f \, d\mu_+, \quad \lambda_-(f) = \int_K f \, d\mu_- \qquad (f \in C_{\mathbb{R}}(K)).$$

Then if $\mu = \mu_+ - \mu_-$, we see that

$$\lambda(f) = \lambda_+(f) - \lambda_-(f) = \int_K f \, d\mu \qquad (f \in C_{\mathbb{R}}(K)).$$

Finally, if $\mathbb{K} = \mathbb{C}$, then we use the same "complexification" trick from the proof of the Hahn-Banach theorem. Namely, let $\lambda \in C_{\mathbb{C}}(K)^*$, and define $\lambda_r, \lambda_i \in C_{\mathbb{R}}(K)^*$ by

$$\lambda_r(f) = \Re\lambda(f), \quad \lambda_i(f) = \Im\lambda(f) \qquad (f \in C_{\mathbb{R}}(K)).$$

These are both clearly $\mathbb{R}$-linear. Notice also that $|\lambda_r(f)| = |\Re\lambda(f)| \leq |\lambda(f)| \leq \|\lambda\|\|f\|_\infty$, so λ_r is bounded; similarly λ_i.

By the real version of the Riesz Representation Theorem, there exist signed measures μ_r and μ_i such that

$$\Re\lambda(f) = \lambda_r(f) = \int_K f \, d\mu_r, \quad \Im\lambda(f) = \lambda_i(f) = \int_K f \, d\mu_i \qquad (f \in C_{\mathbb{R}}(K)).$$

Then let $\mu = \mu_r + i\mu_i$, so for $f \in C_{\mathbb{C}}(K)$,

$$\begin{aligned}
\int_K f \; d\mu &= \int_K f \; d\mu_r + i \int_K f \; d\mu_i \\
&= \int_K \Re(f) \; d\mu_r + i \int_K \Im(f) \; d\mu_r + i \int_K \Re(f) \; d\mu_i - \int_K \Im(f) \; d\mu_i \\
&= \lambda_r(\Re(f)) + i\lambda_r(\Im(f)) + i\lambda_i(\Re(f)) - \lambda_i(\Im(f)) \\
&= \Re\lambda(\Re(f)) + i\Re\lambda(\Im(f)) + i\Im\lambda(\Re(f)) - \Im\lambda(\Im(f)) \\
&= \lambda(\Re(f) + i\Im(f)) = \lambda(f),
\end{aligned}$$

as required. $\square$

Notice that we have not currently proved that $\|\mu\| = \|\lambda\|$ in the case $\mathbb{K} = \mathbb{C}$. See a textbook for this.

Linear Analysis V: Fourier Theory

Updated: 9th December 2009

We shall just develop a tiny bit of Fourier Theory, which is a vast and growing theory. We shall concentrate of the "pure" side the theory. Fourier Theory underpins much of modern signal processing theory: every time you listen to a CD, use a MP3 player, or watch a DVD, you are making use of the Fourier Transform. However, for such "applied" applications, the mathematical formulation usually uses a finite field, for which you do not need analysis, never mind measure theory!

I can strongly recommend the book "Fourier Analysis" by T.W. Körner.

We identify $\mathbb{T}$ with the unit circle, which is the interval $[0,1]$, with the points 0 and 1 identified. Hence we identify $C_{\mathbb{K}}(\mathbb{T})$ with the continuous functions $f : [0,1] \to \mathbb{K}$ such that $f(0) = f(1)$. Equivalently, we identify $C_{\mathbb{K}}(\mathbb{T})$ with the continuous functions $f : \mathbb{R} \to \mathbb{K}$ such that $f(t+n) = f(t)$ for $t \in \mathbb{R}$ and $n \in \mathbb{Z}$ (so $\mathbb{T} = \mathbb{R}/\mathbb{Z}$).

There are some "obvious" functions in $C_{\mathbb{C}}(\mathbb{T})$,

$$\hat{n}(t) = e^{2\pi i n t} = \exp(2\pi i n t) \qquad (n \in \mathbb{Z}, t \in \mathbb{T}).$$

Fourier theory is, basically, interested in decomposition members of $C_{\mathbb{C}}(\mathbb{T})$ into linear combinations of the functions $\{\hat{n} : n \in \mathbb{Z}\}$. As

$$\hat{n}(t) = \cos(2\pi n t) + i \sin(2\pi n t),$$

there are obvious ways to handle real-valued functions. We shall stick to the case $\mathbb{K} = \mathbb{C}$.

The *Fourier Transform* is the map $\mathcal{F} : \mathcal{L}^1([0,1]) \to \ell^\infty(\mathbb{Z})$, defined by

$$\mathcal{F}(f) = \Big(\int_0^1 f(t) e^{2\pi i n t} \ d\mu(t) \Big)_{n \in \mathbb{Z}}.$$

Here $\ell^\infty(\mathbb{Z})$ is just the space of bounded families $(a_n)_{n \in \mathbb{Z}} \subseteq \mathbb{C}$ with the norm $\|(a_n)\| = \sup_n |a_n|$. That is, we just replace $\mathbb{N}$ by $\mathbb{Z}$. Similarly, we can form $\ell^p(\mathbb{Z})$. You need to decide how to add up "sequences" indexed by $\mathbb{Z}$, but there is no problem, as we only care about absolute convergence.

As the set $\{0,1\}$ has zero measure, it is trivial to identify $\mathcal{L}^1([0,1])$ with $\mathcal{L}^1(\mathbb{T})$. I shall continue to swap between these two Banach spaces.

We have the obvious estimate

$$\Big| \int_0^1 f(t) e^{2\pi i n t} \ d\mu(t) \Big| \leq \int_0^1 |f| \ d\mu = \|f\|_1,$$

which shows that $\mathcal{F}$ does map $\mathcal{L}^1([0,1])$ into $\ell^\infty(\mathbb{Z})$. In fact, more is true.

Theorem (Riemann-Lebesgue Lemma): $\mathcal{F}$ maps into the closed subspace $c_0(\mathbb{Z})$ of $\ell^\infty(\mathbb{Z})$.

Proof: Let $0 \leq a < b \leq 1$, let $A = (a,b)$, so that $\chi_A \in \mathcal{L}^1(\mathbb{T})$. Then we calculate:

$$\begin{aligned} \lim_{|n| \to \infty} \Big| \int_{\mathbb{T}} \chi_A \hat{n} \ d\mu \Big| &= \lim_{|n| \to \infty} \Big| \int_a^b e^{2\pi i n t} \ d\mu \Big| = \lim_{|n| \to \infty} \Big| \Big[\frac{e^{2\pi i n t}}{2\pi i n} \Big]_{t=a}^b \ d\mu \Big| \\ &= \lim_{|n| \to \infty} \frac{|e^{2\pi i n b} - e^{2\pi i n a}|}{2\pi |n|} \leq \lim_{|n| \to \infty} \frac{1}{\pi |n|} = 0. \end{aligned}$$

So $\mathcal{F}(\chi_A) \in c_0(\mathbb{Z})$.

We now use the fact that Lebesgue measure is *regular.* If $A \subseteq [0,1]$ is Lebesgue-measurable, then for $\epsilon > 0$ there exists $K \subseteq A$ closed, and $A \subseteq U$ open, with $\mu(U \setminus K) < \epsilon$.

For each $k \in K$, we can find $a < k < b$ such that $V_k = (a,b) \subseteq U$. Thus

$$K \subseteq \bigcup_{k \in V} V_k \subseteq U,$$

so as K is compact, there exists $k_1, \cdots, k_n$ such that $K \subseteq V_{k_1} \cup \cdots \cup V_{k_n}$. Let $V = V_{k_1} \cup \cdots \cup V_{k_n}$, so $K \subseteq V \subseteq U$. If any of the sets V_{k_i} overlap, then we can combine them and still have an open interval. So we can write V as the finite union of *disjoint* open intervals; as $\mathcal{F}$ is linear, it follows that $\mathcal{F}(\chi_V) \in c_0(\mathbb{Z})$. Then notice that

$$\|\chi_A - \chi_V\|_1 = \int_{[0,1]} |\chi_A - \chi_V| \, d\mu = \int_{[0,1]} \chi_{V \setminus A} + \chi_{A \setminus V} \, d\mu = \mu(V \setminus A) + \mu(A \setminus V) < \epsilon.$$

Hence $\|\mathcal{F}(\chi_A) - \mathcal{F}(\chi_V)\| < \epsilon$, as $\|\mathcal{F}\| \leq 1$. As $\epsilon > 0$ was arbitrary, and $c_0(\mathbb{Z})$ is closed, we conclude that $\mathcal{F}(\chi_A) \in c_0(\mathbb{Z})$. (This follows as we can approximate $\mathcal{F}(\chi_A)$ by something (namely $\mathcal{F}(\chi_V)$) in $c_0(\mathbb{Z})$).

As $c_0(\mathbb{Z})$ is a subspace and $\mathcal{F}$ is linear, it follows that $\mathcal{F}(f) \in c_0(\mathbb{Z})$ for any *simple* function $f \in \mathcal{L}^1([0,1])$. In Chapter 3, we showed that simple functions are dense in $\mathcal{L}^1([0,1])$. It follows, again by approximation, that $\mathcal{F}$ does map into $c_0(\mathbb{Z})$. □

Another way to prove this would be to use a uniform continutity argument to show that if $f \in C_{\mathbb{C}}(\mathbb{T})$, then $\mathcal{F}(f) \in c_0(\mathbb{Z})$, and then use that $C_{\mathbb{C}}(\mathbb{T})$ is dense[1] in $\mathcal{L}^1([0,1]) = \mathcal{L}^1(\mathbb{T})$.

The *inverse Fourier Transform*[2] is the map $\mathcal{F}^{-1} : \ell^1(\mathbb{Z}) \to C_{\mathbb{C}}(\mathbb{T})$ defined by

$$\mathcal{F}^{-1}((a_n)_{n \in \mathbb{Z}}) = \Big(\sum_{n \in \mathbb{Z}} a_n e^{-2\pi i n t} \Big)_{t \in \mathbb{T}}.$$

Notice that the sum is absolutely convergent in the Banach space $C_{\mathbb{C}}(\mathbb{T})$, and so certainly converges.

Lemma: If $a = (a_n)_{n \in \mathbb{Z}} \in \ell^1(\mathbb{Z})$, then $\mathcal{F}(\mathcal{F}^{-1}(a)) = (a_n)$, where we treat (a_n) as a sequence in $c_0(\mathbb{Z})$. In particular, $\mathcal{F}^{-1}$ is injective.

Proof: Notice that for $n, m \in \mathbb{Z}$,

$$\int_{[0,1]} e^{2\pi i n t} e^{-2\pi i m t} \, d\mu(t) = \int_0^1 e^{2\pi i t(n-m)} \, dt = 1 \text{ if } n = m,$$

$$\text{otherwise, } = \Big[\frac{e^{2\pi i t(n-m)}}{2\pi i (n-m)} \Big]_{t=0}^1 = 0.$$

Let $f_n \in C_{\mathbb{C}}(\mathbb{T})$ be defined by

$$f_n(t) = \sum_{k=-n}^{n} a_k e^{-2\pi i k t} \qquad (t \in \mathbb{T}).$$

Then let $f = \mathcal{F}^{-1}(a)$, so that

$$\|f - f_n\|_\infty = \sup_{t \in \mathbb{T}} \Big| \sum_{|k| > n} a_k e^{-2\pi i k t} \Big| \leq \sum_{|k|>n} |a_k| \to 0,$$

[1] We haven't proved this, but it's on the example sheet, and isn't very hard.

[2] Notice that, in the set theoretic sense, $\mathcal{F}$ and $\mathcal{F}^{-1}$ are *not* inverses to each other, as they have different domains and codomains. We clarify this later.

as $n \to \infty$. As $\mathcal{F}$ is bounded,

$$\mathcal{F}(f) = \lim_n \mathcal{F}(f_n) = \lim_n \Big(\int_{[0,1]} \sum_{k=-n}^{n} a_k e^{-2\pi i k t} e^{2\pi i m t} \ d\mu(t) \Big)_{m\in\mathbb{Z}}$$
$$= \lim_n \Big(\sum_{k=-n}^{n} a_k \int_0^1 e^{-2\pi i k t} e^{2\pi i m t} \ dt \Big)_{m\in\mathbb{Z}} = (a_m)_{m\in\mathbb{Z}}.$$

As $\mathcal{F}^{-1}$ has a left inverse, it must be injective. $\square$

So the "inverse Fourier Transform" is, in a loose sense, the inverse to the "Fourier Transform", as long as we're working in the correct spaces.

Classically, there was a lot of interest in the pointwise limit convergence of Fourier Series. Namely, if $f : \mathbb{T} \to \mathbb{C}$ is continuous, if we let $\mathcal{F}(f) = (\hat{f}(n))_{n\in\mathbb{Z}} \in c_0(\mathbb{Z})$, and if we define

$$S_n(f,t) = \sum_{|k|\leq n} \hat{f}(k) e^{-2\pi i k t} \qquad (n \in \mathbb{N}, t \in \mathbb{T}),$$

then when is it true that $\lim_n S_n(f,t) = f(t)$?

Claim: There exists a continuous function $f : \mathbb{T} \to \mathbb{C}$ such that $\limsup_n S_n(f,0) = \infty$.

So pointwise limits can be badly behaved. Remarkably, if we "tweak" the convergence method, then we *can* always recover f.

Theorem (Fejer's Theorem): Let $f : \mathbb{T} \to \mathbb{C}$ be a continuous function, and for $n \in \mathbb{N}$, define

$$\sigma_n(t) = \frac{1}{n+1} \sum_{k=0}^{n} S_k(f,t) = \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} \hat{f}(k) e^{-2\pi i k t} \qquad (t \in \mathbb{T}).$$

Then $\sigma_n \to f$ in $C_{\mathbb{C}}(\mathbb{T})$ (that is, uniform convergence).

Proof: We calculate:

$$\sigma_n(t) = \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} \hat{f}(k) e^{-2\pi i k t}$$
$$= \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} \int_{[0,1]} f(s) e^{2\pi i k s} \ d\mu(s) e^{-2\pi i k t}$$
$$= \int_{[0,1]} f(s) \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} e^{2\pi i k (s-t)} \ d\mu(s)$$
$$= \int_{[0,1]} f(r+t) \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} e^{2\pi i k r} \ d\mu(r).$$

Here we interpret f and $e^{2\pi i k \cdot}$ as periodic functions, so we don't need to change the interval which we integrate over. So if we define

$$K_n(r) = \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} e^{2\pi i k r} \qquad (r \in \mathbb{T}),$$

then studying K_n is obviously a good idea!

We first show by direct calculation that

$$K_n(0) = n+1, \quad K_n(s) = \frac{1}{n+1} \Big(\frac{\sin(\pi s(n+1))}{\sin(\pi s)} \Big)^2 \qquad (s \neq 0).$$

Once we have this desciption for K_n, it is easy to show:

1. $K_n \geq 0$ for all n;
2. For any $\delta > 0$, on the interval $(\delta, 1-\delta)$, we have that $K_n \to 0$ uniformly;
3. $\int_{[0,1]} K_n \, d\mu = 1$ for all n.

We can now finish the proof. We see that, by (3),

$$|\sigma_n(t) - f(t)| = \Big| \int_{[0,1]} \big(f(r+t) - f(t)\big) K_n(r) \, d\mu(r) \Big|$$

Motivated by (2), and using (1), this is

$$\leq \int_{[0,\delta]} |f(r+t) - f(t)| K_n(r) \, d\mu(r) + \int_{[1-\delta,1]} |f(r+t) - f(t)| K_n(r) \, d\mu(r) \\ + \int_{(\delta,1-\delta)} |f(r+t) - f(t)| K_n(r) \, d\mu(r).$$

As f is uniformly continuous, for $\epsilon > 0$, we can find $\delta > 0$ such that $|f(r+t) - f(t)| < \epsilon$ if $|r| < \delta$. As f is periodic, this also holds if $|1-r| < \delta$. By (2), if n is sufficiently large, then $\sup_{\delta < r < 1-\delta} |K_n(r)| < \epsilon$. So we find that

$$\leq \epsilon \int_{[0,\delta]} K_n(r) \, d\mu(r) + \epsilon \int_{[1-\delta,1]} K_n(r) \, d\mu(r) + 2\|f\|_\infty \epsilon < 2\epsilon(1 + \|f\|_\infty),$$

using (3) again. So $\|\sigma_n - f\|_\infty \to 0$, as required. $\square$

Corollary: If we restrict $\mathcal{F}$ to $C_{\mathbb{C}}(\mathbb{T})$, then $\mathcal{F}$ is injective. In particular, if $f \in C_{\mathbb{C}}(\mathbb{T})$ is such that $\mathcal{F}(f) \in \ell^1(\mathbb{Z})$, then $\mathcal{F}^{-1}\mathcal{F}(f) = f$.

Proof: If $f, g \in C_{\mathbb{C}}(\mathbb{T})$ are such that $\mathcal{F}(f) = \mathcal{F}(g)$, then let $h = f - g$, so $\mathcal{F}(h) = 0$. By Fejer's Theorem, we can reconstruct h from $\mathcal{F}(h)$, but as $\mathcal{F}(h) = 0$, we reconstruct the 0 function, so $h = 0$. So $f = g$.

If $\mathcal{F}(f) = a \in \ell^1(\mathbb{Z})$, then $\mathcal{F}^{-1}(a)$ is defined, and equals, say, $g \in C_{\mathbb{C}}(\mathbb{T})$. Then by a previous result, $\mathcal{F}(g) = \mathcal{F}\mathcal{F}^{-1}(a) = a = \mathcal{F}(f)$. So $\mathcal{F}(f) = \mathcal{F}(g)$, so $f = g$. $\square$

The following is slightly tricky, but is a nice application of some of the ideas which we have seen in the course.

Corollary: $\mathcal{F} : \mathcal{L}^1([0,1]) \to c_0(\mathbb{Z})$ is injective.

Proof: We exploit "duality". Let $f \in \mathcal{L}^1([0,1])$ and let $a = (a_n) \in \ell^1(\mathbb{Z})$, so that

$$\int_{[0,1]} f \, \mathcal{F}^{-1}(a) \, d\mu = \int_{[0,1]} f(s) \sum_{n \in \mathbb{Z}} a_n e^{-2\pi ins} \, d\mu(s) = \sum_{n \in \mathbb{Z}} a_n \int_{[0,1]} f(s) e^{-2\pi ins} \, d\mu(s).$$

As $(a_n) \in \ell^1(\mathbb{Z})$, the sum is absolutely convergent, and so we can swap the sum and integrals (this *does not* need Monotone Convergence, or something similar: it's a simple estimate). Thus

$$\int_{[0,1]} f \, \mathcal{F}^{-1}(a) \, d\mu = \sum_{n \in \mathbb{Z}} a_n \mathcal{F}(f)(-n).$$

Suppose now that $\mathcal{F}(f) = 0$. So $\int f \mathcal{F}^{-1}(a) \, d\mu = 0$ for any $a \in \ell^1(\mathbb{Z})$.

Let λ be the measure $f\mu$, defined on the Borel sigma algebra $\mathcal{B}([0,1])$. That is,

$$\lambda(A) = \int_{[0,1]} \chi_A f \, d\mu \qquad (A \in \mathcal{B}([0,1])).$$

We know that the Lebesgue measurable sets contain the open sets, and hence contain all the Borel sets. So this integral is defined, as $\chi_A f$ is measurable for μ, and integrable, as $f \in \mathcal{L}^1(\mu)$. A Dominated Convergence theorem argument shows that λ is countably additive, and so is a measure. It is actually a *regular* measure.[3] So λ induces a member of $C_{\mathbb{C}}([0,1])^*$, say ϕ_λ, given by

$$\phi_\lambda(g) = \int_{[0,1]} g \, d\lambda = \int_{[0,1]} gf \, d\mu.$$

The final equality needs proof, but isn't very hard (approximate by simple functions, then apply a convergence theorem!)

We have proved that if g is in the image of $\mathcal{F}^{-1}$, then $\phi_\lambda(g) = 0$. For arbitrary $g \in C_{\mathbb{C}}(\mathbb{T})$, form σ_n as in Fejer's Theorem, so that $\sigma_n \to g$ in $C_{\mathbb{C}}(\mathbb{T})$. Each σ_n is a *finite* sum of functions of the form $e^{2\pi int}$, and so σ_n is in the image of $\mathcal{F}^{-1}$. Thus

$$\phi_\lambda(g) = \lim_n \phi_\lambda(\sigma_n) = 0.$$

As this is true for all $g \in C_{\mathbb{C}}(\mathbb{T})$, it follows that $\phi_\lambda = 0$. By Riesz Representation, ϕ_λ is induced by a *unique* finite regular Borel measure. As λ is regular, this measure must be λ. But as $\phi_\lambda = 0$, it follows that $\lambda = 0$. But this can only happen if $f = 0$. So $\mathcal{F}$ is injective. □

We now turn our attention to the Banach space $\mathcal{L}^2(\mathbb{T})$.

Lemma: Let $(a_n) \in \ell^2(\mathbb{Z})$ be such that $a_n = 0$ for all but finitely many n. Then $\|\mathcal{F}^{-1}((a_n))\|_2 = \|(a_n)\|_2$.

Proof: Notice that under the conditions on (a_n), we have that $(a_n) \in \ell^1(\mathbb{Z})$, and so $\mathcal{F}^{-1}((a_n))$ is defined. Then we calculate

$$\begin{aligned}
\|\mathcal{F}^{-1}((a_n))\|_2^2 &= \int_{[0,1]} \Big| \sum_n a_n e^{-2\pi int} \Big|^2 \, d\mu(t) \\
&= \int_{[0,1]} \Big(\sum_n a_n e^{-2\pi int} \Big) \Big(\sum_m \overline{a_m} e^{2\pi imt} \Big) \, d\mu(t) \\
&= \sum_{n,m} a_n \overline{a_m} \int_0^1 e^{2\pi i(m-n)t} \, d\mu(t) = \sum_n a_n \overline{a_n} = \|(a_n)\|_2^2.
\end{aligned}$$

We can do the manipulations, as all the sums are really just finite sums. □

We now prove an abstract result.

Proposition: Let E and F be Banach spaces, let $X \subseteq E$ be a dense subspace, and let $T : X \to F$ be an isometry which has dense range. Then T extends uniquely to an isometric isomorphism $\tilde{T} : E \to F$.

Proof: Let $x \in E$, so we can find a sequence (x_n) in X with $x_n \to x$. Then (x_n) is Cauchy, but as T is an isometry,

$$\|T(x_n) - T(x_m)\| = \|T(x_n - x_m)\| = \|x_n - x_m\|,$$

and so we see that $(T(x_n))$ is a Cauchy sequence in F. As F is Banach, there exists a limit point, say $\tilde{T}(x)$. So $T(x_n) \to \tilde{T}(x)$.

[3] Strictly non-examinable! Let $A \in \mathcal{B}([0,1])$. As μ is regular, for each n, we can find an open set U_n and a closed set K_n with $K_n \subseteq A \subseteq U_n$ with $\mu(U_n \setminus K_n) < 1/n$. For each n, let $V_n = U_1 \cap \cdots \cap U_n$ and $C_n = K_1 \cup \cdots \cup K_n$, so V_n is open, C_n is closed, $C_n \subseteq A \subseteq V_n$, and still $\mu(V_n \setminus C_n) < 1/n$. Now $V_1 \supseteq V_2 \supseteq V_3 \supseteq \cdots$ and $C_1 \subseteq C_2 \subseteq C_3 \subseteq \cdots$, so also $V_1 \setminus C_1 \supseteq V_2 \setminus C_2 \supseteq \cdots$. Then a simple calculation shows that $\chi_{V_n \setminus C_n} \to 0$ almost everywhere. So also $|f|\chi_{V_n \setminus C_n} \to 0$ almost everywhere. As this sequence is dominated by $|f|$, by Dominated Convergence, $\int |f|\chi_{V_n \setminus C_n} \to 0$. This is enough to show that $|\lambda|$ is regular, which shows that λ itself is regular: see the final example sheet.

If we pick another sequence (y_n) with $y_n \to x$, then $(x_n - y_n)$ converges to $x - x = 0$, and so also $T(x_n) - T(y_n) \to 0$, so $T(y_n) \to \tilde{T}(x)$. So $\tilde{T} : E \to F$ is well-defined.

If $x_n \to x$ and $y_n \to y$, for $t \in \mathbb{K}$, we have that $x_n + ty_n \to x + ty$, and so

$$\tilde{T}(x + ty) = \lim_n T(x_n + ty_n) = \lim_n T(x_n) + \lim_n tT(y_n) = \tilde{T}(x) + t\tilde{T}(y).$$

So $\tilde{T}$ is linear. Also $\|\tilde{T}(x)\| = \lim_n \|T(x_n)\| = \lim_n \|x_n\| = \|x\|$, so we conclude that $\tilde{T}$ is an isometry.

If S is another continuous extension of T, then

$$S(x) = \lim_n S(x_n) = \lim_n T(x_n) = \tilde{T}(x),$$

so $S = \tilde{T}$. Hence $\tilde{T}$ is unique.

Finally, we show that $\tilde{T}$ is a surjection (it is automatically an injection, as it is an isometry). Let $y \in F$, so as $T(X)$ is dense in Y by hypothesis, we can find (x_n) in X with $T(x_n) \to y$. Then $(T(x_n))$ is Cauchy in Y, so as T is an isometry, (x_n) is Cauchy in X, and hence converges to $x \in E$, say. Then $\tilde{T}(x) = \lim_n T(x_n) = y$, so $\tilde{T}$ is surjective.$\square$

Proposition: We have that $\mathcal{F}^{-1}$ extends to an isometry $\ell^2(\mathbb{Z}) \to \mathcal{L}^2([0,1])$.

Proof: We have that $\mathcal{F}^{-1}$ is an isometry on a dense subspace of $\ell^2(\mathbb{Z})$, namely the space of finite sequences. Its image is just the linear span of $\{\hat{n} : n \in \mathbb{Z}\}$. By Fejer, this space is dense in $C_{\mathbb{C}}(\mathbb{T})$, which is dense in $\mathcal{L}^2(\mathbb{T})$. These norms are different, so let's be careful. Given $f \in \mathcal{L}^2([0,1])$ and $\epsilon > 0$, we can find[4] $g \in C_{\mathbb{C}}(\mathbb{T})$ with $\|f - g\|_2 < \epsilon$. Then we can find σ, a finite combination of $\{\hat{n} : n \in \mathbb{Z}\}$, with $\|g - \sigma\|_\infty < \epsilon$. A simple calculation (use Holder!) shows that $\|g - \sigma\|_2 < \epsilon$. So $\|f - \sigma\|_2 < 2\epsilon$.

So all the conditions of the previous result are satisfied, and so we conclude that $\mathcal{F}^{-1}$ has a unique extension to an isometric isomorphism $\ell^2(\mathbb{Z}) \to \mathcal{L}^2([0,1])$. $\square$

Proposition: $\mathcal{F}$, as defined by the formula at the start of this chapter, actually makes sense as a map $\mathcal{F} : \mathcal{L}^2([0,1]) \to \ell^\infty(\mathbb{Z})$. Then really $\mathcal{F}$ is an isometric isomorphism $\mathcal{L}^2([0,1]) \to \ell^2(\mathbb{Z})$, which is the inverse of (the extension of) $\mathcal{F}^{-1}$.

Proof: Notice that actually $\mathcal{L}^2([0,1]) \subseteq \mathcal{L}^1([0,1])$, for if $f \in \mathcal{L}^2([0,1])$, then by Holder,

$$\int_{[0,1]} |f| \; d\mu \leq \Big(\int_{[0,1]} |f|^2 \; d\mu \Big)^{1/2} \Big(\int_{[0,1]} 1 \; d\mu \Big)^{1/2} = \|f\|_2.$$

So $\mathcal{F}$ is already defined upon $\mathcal{L}^2([0,1])$.

Again, let X be the linear span of the functions $(e^{2\pi int})_{t\in[0,1]}$, for $n \in \mathbb{Z}$. In the previous proof, we showed that X is a dense subspace of $\mathcal{L}^2([0,1])$. For $f \in X$ we have that $f = \mathcal{F}(a)$ for some sequence $a = (a_n) \in \ell^2(\mathbb{Z})$ which is zero except in finitely many places. The lemma above shows that

$$\|a\|_2 = \|f\|_2 \implies \|\mathcal{F}(f)\|_2 = \|a\|_2 = \|f\|_2.$$

So $\mathcal{F}$ is an isometry $X \to \ell^2(\mathbb{Z})$, with image the finite sequences (which is a dense subspace of $\ell^2(\mathbb{Z})$. So, again by our abstract result, $\mathcal{F}$ extends uniquely to an isometric isomorphism. As this extension is unique, and $\mathcal{F}$ is already defined on all of $\mathcal{L}^2([0,1])$, then extension must just be $\mathcal{F}$. $\square$

Finally, turning attention back to pointwise convergence, we have the following theorems, stated without proof.

[4]For all the details, see the example sheet

Theorem (Kolmogorov): There exists a Lebesgue integrable function $f : \mathbb{T} \to \mathbb{C}$ such that

$$\limsup_n \left|S_n(f,t)\right| = \infty$$

for all $t \in \mathbb{T}$.

So $S_n(f,t)$ can fail, spectacularly, to converge. However, if f is continuous, remarkably, we have the following:

Theorem (Carleson): Let $f \in C_{\mathbb{C}}(\mathbb{T})$. Then there exists a null set $E \subseteq \mathbb{T}$ such that, if $t \notin E$, then $S_n(f,t) \to f(t)$.

Proof: Hard![5]

An easier argument establishes the converse.

Theorem (Kahane and Katznelson): If $E \subseteq \mathbb{T}$ is a null set, then there exists $f \in C_{\mathbb{C}}(\mathbb{T})$ with $\limsup_n |S_n(f,t)| = \infty$ for $t \in E$.

[5]To the extent that this is widely considered to be *the* hardest result is Fourier analysis: I don't know enough to judge if this is fair or not!