1 L^p Spaces and Banach Spaces

In this work the assumption of quadratic integrability will be replaced by the integrability of $|f(x)|^p$. The analysis of these function classes will shed a particular light on the real and apparent advantages of the exponent 2; one can also expect that it will provide essential material for an axiomatic study of function spaces.

F. Riesz, 1910

At present I propose above all to gather results about linear operators defined in certain general spaces, notably those that will here be called *spaces of type (B)...*

S. Banach, 1932

Function spaces, in particular L^p spaces, play a central role in many questions in analysis. The special importance of L^p spaces may be said to derive from the fact that they offer a partial but useful generalization of the fundamental L^2 space of square integrable functions.

In order of logical simplicity, the space L^1 comes first since it occurs already in the description of functions integrable in the Lebesgue sense. Connected to it via duality is the L^∞ space of bounded functions, whose supremum norm carries over from the more familiar space of continuous functions. Of independent interest is the L^2 space, whose origins are tied up with basic issues in Fourier analysis. The intermediate L^p spaces are in this sense an artifice, although of a most inspired and fortuitous kind. That this is the case will be illustrated by results in the next and succeeding chapters.

In this chapter we will concentrate on the basic structural facts about the L^p spaces. Here part of the theory, in particular the study of their linear functionals, is best formulated in the more general context of Banach spaces. An incidental benefit of this more abstract view-point is that it leads us to the surprising discovery of a finitely additive measure on *all* subsets, consistent with Lebesgue measure.

1 L^p spaces

Throughout this chapter $(X, \mathcal{F}, \mu)$ denotes a σ-finite measure space: X denotes the underlying space, $\mathcal{F}$ the σ-algebra of measurable sets, and μ the measure. If $1 \leq p < \infty$, the space $L^p(X, \mathcal{F}, \mu)$ consists of all complex-valued measurable functions on X that satisfy

$$\int_X |f(x)|^p \, d\mu(x) < \infty. \tag{1}$$

To simplify the notation, we write $L^p(X, \mu)$, or $L^p(X)$, or simply L^p when the underlying measure space has been specified. Then, if $f \in L^p(X, \mathcal{F}, \mu)$ we define the L^p **norm** of f by

$$\|f\|_{L^p(X,\mathcal{F},\mu)} = \left(\int_X |f(x)|^p \, d\mu(x) \right)^{1/p}.$$

We also abbreviate this to $\|f\|_{L^p(X)}$, $\|f\|_{L^p}$, or $\|f\|_p$.

When $p = 1$ the space $L^1(X, \mathcal{F}, \mu)$ consists of all integrable functions on X, and we have shown in Chapter 6 of Book III, that L^1 together with $\|\cdot\|_{L^1}$ is a complete normed vector space. Also, the case $p = 2$ warrants special attention: it is a Hilbert space.

We note here that we encounter the same technical point that we already discussed in Book III. The problem is that $\|f\|_{L^p} = 0$ does not imply that $f = 0$, but merely $f = 0$ almost everywhere (for the measure μ). Therefore, the precise definition of L^p requires introducing the equivalence relation, in which f and g are equivalent if $f = g$ a.e. Then, L^p consists of all equivalence classes of functions which satisfy (1). However, in practice there is little risk of error by thinking of elements in L^p as functions rather than equivalence classes of functions.

The following are some common examples of L^p spaces.

(a) The case $X = \mathbb{R}^d$ and μ equals Lebesgue measure is often used in practice. There, we have

$$\|f\|_{L^p} = \left(\int_{\mathbb{R}^d} |f(x)|^p \, dx \right)^{1/p}.$$

(b) Also, one can take $X = \mathbb{Z}$, and μ equal to the counting measure. Then, we get the "discrete" version of the L^p spaces. Measurable functions are simply sequences $f = \{a_n\}_{n \in \mathbb{Z}}$ of complex numbers, and

$$\|f\|_{L^p} = \left(\sum_{n=-\infty}^{\infty} |a_n|^p \right)^{1/p}.$$

When $p = 2$, we recover the familiar sequence space $\ell^2(\mathbb{Z})$.

The spaces L^p are examples of normed vector spaces. The basic property satisfied by the norm is the triangle inequality, which we shall prove shortly.

The range of p which is of interest in most applications is $1 \leq p < \infty$, and later also $p = \infty$. There are at least two reasons why we restrict our attention to these values of p: when $0 < p < 1$, the function $\|\cdot\|_{L^p}$ does not satisfy the triangle inequality, and moreover, for such p, the space L^p has no non-trivial bounded linear functionals.[1] (See Exercise 2.)

When $p = 1$ the norm $\|\cdot\|_{L^1}$ satisfies the triangle inequality, and L^1 is a complete normed vector space. When $p = 2$, this result continues to hold, although one needs the Cauchy-Schwarz inequality to prove it. In the same way, for $1 \leq p < \infty$ the proof of the triangle inequality relies on a generalized version of the Cauchy-Schwarz inequality. This is Hölder's inequality, which is also the key in the duality of the L^p spaces, as we will see in Section 4.

1.1 The Hölder and Minkowski inequalities

If the two exponents p and q satisfy $1 \leq p, q \leq \infty$, and the relation

$$\frac{1}{p} + \frac{1}{q} = 1$$

holds, we say that p and q are **conjugate** or **dual** exponents. Here, we use the convention $1/\infty = 0$. Later, we shall sometimes use p' to denote the conjugate exponent of p. Note that $p = 2$ is self-dual, that is, $p = q = 2$; also $p = 1, \infty$ corresponds to $q = \infty, 1$ respectively.

Theorem 1.1 (Hölder) *Suppose $1 < p < \infty$ and $1 < q < \infty$ are conjugate exponents. If $f \in L^p$ and $g \in L^q$, then $fg \in L^1$ and*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Note. Once we have defined L^∞ (see Section 2) the corresponding inequality for the exponents 1 and ∞ will be seen to be essentially trivial.

---

[1] We will define what we mean by a bounded linear functional later in the chapter.

The proof of the theorem relies on a simple generalized form of the arithmetic-geometric mean inequality: if $A, B \geq 0$, and $0 \leq \theta \leq 1$, then

$$A^\theta B^{1-\theta} \leq \theta A + (1-\theta)B. \tag{2}$$

Note that when $\theta = 1/2$, the inequality (2) states the familiar fact that the geometric mean of two numbers is majorized by their arithmetic mean.

To establish (2), we observe first that we may assume $B \neq 0$, and replacing A by AB, we see that it suffices to prove that $A^\theta \leq \theta A + (1-\theta)$. If we let $f(x) = x^\theta - \theta x - (1-\theta)$, then $f'(x) = \theta(x^{\theta-1} - 1)$. Thus $f(x)$ increases when $0 \leq x \leq 1$ and decreases when $1 \leq x$, and we see that the continuous function f attains a maximum at $x = 1$, where $f(1) = 0$. Therefore $f(A) \leq 0$, as desired.

To prove Hölder's inequality we argue as follows. If either $\|f\|_{L^p} = 0$ or $\|f\|_{L^q} = 0$, then $fg = 0$ a.e. and the inequality is obviously verified. Therefore, we may assume that neither of these norms vanish, and after replacing f by $f/\|f\|_{L^p}$ and g by $g/\|g\|_{L^q}$, we may further assume that $\|f\|_{L^p} = \|g\|_{L^q} = 1$. We now need to prove that $\|fg\|_{L^1} \leq 1$.

If we set $A = |f(x)|^p$, $B = |g(x)|^q$, and $\theta = 1/p$ so that $1 - \theta = 1/q$, then (2) gives

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

Integrating this inequality yields $\|fg\|_{L^1} \leq 1$, and the proof of the Hölder inequality is complete.

For the case when the equality $\|fg\|_{L^1} = \|f\|_{L^p}\|g\|_{L^q}$ holds, see Exercise 3.

We are now ready to prove the triangle inequality for the L^p norm.

Theorem 1.2 (Minkowski) *If* $1 \leq p < \infty$ *and* $f, g \in L^p$, *then* $f + g \in L^p$ *and* $\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$.

Proof. The case $p = 1$ is obtained by integrating $|f(x) + g(x)| \leq |f(x)| + |g(x)|$. When $p > 1$, we may begin by verifying that $f + g \in L^p$, when both f and g belong to L^p. Indeed,

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

as can be seen by considering separately the cases $|f(x)| \leq |g(x)|$ and $|g(x)| \leq |f(x)|$. Next we note that

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

If q denotes the conjugate exponent of p, then $(p-1)q = p$, so we see that $(f+g)^{p-1}$ belongs to L^q, and therefore Hölder's inequality applied to the two terms on the right-hand side of the above inequality gives

$$(3) \qquad \|f+g\|_{L^p}^p \leq \|f\|_{L^p}\|(f+g)^{p-1}\|_{L^q} + \|g\|_{L^p}\|(f+g)^{p-1}\|_{L^q}.$$

However, using once again $(p-1)q = p$, we get

$$\|(f+g)^{p-1}\|_{L^q} = \|f+g\|_{L^p}^{p/q}.$$

From (3), since $p - p/q = 1$, and because we may suppose that $\|f+g\|_{L^p} > 0$, we find

$$\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p},$$

so the proof is finished.

1.2 Completeness of L^p

The triangle inequality makes L^p into a metric space with distance $d(f,g) = \|f-g\|_{L^p}$. The basic analytic fact is that L^p is **complete** in the sense that every Cauchy sequence in the norm $\|\cdot\|_{L^p}$ converges to an element *in* L^p.

Taking limits is a necessity in many problems, and the L^p spaces would be of little use if they were not complete. Fortunately, like L^1 and L^2, the general L^p space does satisfy this desirable property.

Theorem 1.3 *The space $L^p(X, \mathcal{F}, \mu)$ is complete in the norm $\|\cdot\|_{L^p}$.*

Proof. The argument is essentially the same as for L^1 (or L^2); see Section 2, Chapter 2 and Section 1, Chapter 4 in Book III. Let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence in L^p, and consider a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ of $\{f_n\}$ with the following property $\|f_{n_{k+1}} - f_{n_k}\|_{L^p} \leq 2^{-k}$ for all $k \geq 1$. We now consider the series whose convergence will be seen below

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|,$$

and the corresponding partial sums

$$S_K(f)(x) = f_{n_1}(x) + \sum_{k=1}^{K} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$S_K(g)(x) = |f_{n_1}(x)| + \sum_{k=1}^{K} |f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

The triangle inequality for L^p implies

$$\begin{aligned} \|S_K(g)\|_{L^p} &\leq \|f_{n_1}\|_{L^p} + \sum_{k=1}^{K} \|f_{n_{k+1}} - f_{n_k}\|_{L^p} \\ &\leq \|f_{n_1}\|_{L^p} + \sum_{k=1}^{K} 2^{-k}. \end{aligned}$$

Letting K tend to infinity, and applying the monotone convergence theorem proves that $\int g^p < \infty$, and therefore the series defining g, and hence the series defining f converges almost everywhere, and $f \in L^p$.

We now show that f is the desired limit of the sequence $\{f_n\}$. Since (by construction of the telescopic series) the $(K-1)^{\text{th}}$ partial sum of this series is precisely f_{n_K}, we find that

$$f_{n_K}(x) \to f(x) \qquad \text{a.e. } x.$$

To prove that $f_{n_K} \to f$ in L^p as well, we first observe that

$$\begin{aligned} |f(x) - S_K(f)(x)|^p &\leq [2\max(|f(x)|, |S_K(f)(x)|)]^p \\ &\leq 2^p|f(x)|^p + 2^p|S_K(f)(x)|^p \\ &\leq 2^{p+1}|g(x)|^p, \end{aligned}$$

for all K. Then, we may apply the dominated convergence theorem to get $\|f_{n_K} - f\|_{L^p} \to 0$ as K tends to infinity.

Finally, the last step of the proof consists of recalling that $\{f_n\}$ is Cauchy. Given $\epsilon > 0$, there exists N so that for all $n, m > N$ we have $\|f_n - f_m\|_{L^p} < \epsilon/2$. If n_K is chosen so that $n_K > N$, and $\|f_{n_K} - f\|_{L^p} < \epsilon/2$, then the triangle inequality implies

$$\|f_n - f\|_{L^p} \leq \|f_n - f_{n_K}\|_{L^p} + \|f_{n_K} - f\|_{L^p} < \epsilon$$

whenever $n > N$. This concludes the proof of the theorem.

1.3 Further remarks

We begin by looking at some possible inclusion relations between the various L^p spaces. The matter is simple if the underlying space has finite measure.

Proposition 1.4 *If X has finite positive measure, and $p_0 \leq p_1$, then $L^{p_1}(X) \subset L^{p_0}(X)$ and*

$$\frac{1}{\mu(X)^{1/p_0}} \|f\|_{L^{p_0}} \leq \frac{1}{\mu(X)^{1/p_1}} \|f\|_{L^{p_1}}.$$

We may assume that $p_1 > p_0$. Suppose $f \in L^{p_1}$, and set $F = |f|^{p_0}$, $G = 1$, $p = p_1/p_0 > 1$, and $1/p + 1/q = 1$, in Hölder's inequality applied to F and G. This yields

$$\|f\|_{L^{p_0}}^{p_0} \leq \left(\int |f|^{p_1}\right)^{p_0/p_1} \cdot \mu(X)^{1-p_0/p_1}.$$

In particular, we find that $\|f\|_{L^{p_0}} < \infty$. Moreover, by taking the p_0^{th} root of both sides of the above equation, we find that the inequality in the proposition holds.

However, as is easily seen, such inclusion does not hold when X has infinite measure. (See Exercise 1). Yet, in an interesting special case the opposite inclusion does hold.

Proposition 1.5 *If $X = \mathbb{Z}$ is equipped with counting measure, then the reverse inclusion holds, namely $L^{p_0}(\mathbb{Z}) \subset L^{p_1}(\mathbb{Z})$ if $p_0 \leq p_1$. Moreover, $\|f\|_{L^{p_1}} \leq \|f\|_{L^{p_0}}$.*

Indeed, if $f = \{f(n)\}_{n\in\mathbb{Z}}$, then $\sum |f(n)|^{p_0} = \|f\|_{L^{p_0}}^{p_0}$, and $\sup_n |f(n)| \leq \|f\|_{L^{p_0}}$. However

$$\begin{aligned} |f(n)|^{p_1} = & \quad |f(n)|^{p_0}|f(n)|^{p_1-p_0} \\ & \leq (\sup_n |f(n)|)^{p_1-p_0}\|f\|_{L^{p_0}}^{p_0} \\ & \leq \|f\|_{L^{p_0}}^{p_1}. \end{aligned}$$

Thus $\|f\|_{L^{p_1}} \leq \|f\|_{L^{p_0}}$.

2 The case $p = \infty$

Finally, we also consider the limiting case $p = \infty$. The space L^∞ will be defined as all functions that are "essentially bounded" in the following sense. We take the space $L^\infty(X, \mathcal{F}, \mu)$ to consist of all (equivalence classes of) measurable functions on X, so that there exists a positive number $0 < M < \infty$, with

$$|f(x)| \leq M \qquad \text{a.e. } x.$$

Then, we define $\|f\|_{L^\infty(X,\mathcal{F},\mu)}$ to be the infimum of all possible values M satisfying the above inequality. The quantity $\|f\|_{L^\infty}$ is sometimes called the **essential-supremum** of f.

We note that with this definition, we have $|f(x)| \leq \|f\|_{L^\infty}$ for a.e. x. Indeed, if $E = \{x : |f(x)| > \|f\|_{L^\infty}\}$, and $E_n = \{x : |f(x)| > \|f\|_{L^\infty} + 1/n\}$, then we have $\mu(E_n) = 0$, and $E = \bigcup E_n$, hence $\mu(E) = 0$.

Theorem 2.1 *The vector space L^∞ equipped with $\|\cdot\|_{L^\infty}$ is a complete vector space.*

This assertion is easy to verify and is left to the reader. Moreover, Hölder's inequality continues to hold for values of p and q in the larger range $1 \leq p, q \leq \infty$, once we take $p = 1$ and $q = \infty$ as conjugate exponents, as we mentioned before.

The fact that L^∞ is a limiting case of L^p when p tends to ∞ can be understood as follows.

Proposition 2.2 *Suppose $f \in L^\infty$ is supported on a set of finite measure. Then $f \in L^p$ for all $p < \infty$, and*

$$\|f\|_{L^p} \to \|f\|_{L^\infty} \qquad \text{as } p \to \infty.$$

Proof. Let E be a measurable subset of X with $\mu(E) < \infty$, and so that f vanishes in the complement of E. If $\mu(E) = 0$, then $\|f\|_{L^\infty} = \|f\|_{L^p} = 0$ and there is nothing to prove. Otherwise

$$\|f\|_{L^p} = \left(\int_E |f(x)|^p \, d\mu\right)^{1/p} \leq \left(\int_E \|f\|_{L^\infty}^p \, d\mu\right)^{1/p} \leq \|f\|_{L^\infty}\mu(E)^{1/p}.$$

Since $\mu(E)^{1/p} \to 1$ as $p \to \infty$, we find that $\limsup_{p\to\infty} \|f\|_{L^p} \leq \|f\|_{L^\infty}$.

On the other hand, given $\epsilon > 0$, we have

$$\mu(\{x : \ |f(x)| \geq \|f\|_{L^\infty} - \epsilon\}) \geq \delta \qquad \text{for some } \delta > 0,$$

hence

$$\int_X |f|^p \, d\mu \geq \delta(\|f\|_{L^\infty} - \epsilon)^p.$$

Therefore $\liminf_{p\to\infty} \|f\|_{L^p} \geq \|f\|_{L^\infty} - \epsilon$, and since ϵ is arbitrary, we have $\liminf_{p\to\infty} \|f\|_{L^p} \geq \|f\|_{L^\infty}$. Hence the limit $\lim_{p\to\infty} \|f\|_{L^p}$ exists, and equals $\|f\|_{L^\infty}$.

3 Banach spaces

We introduce here a general notion which encompasses the L^p spaces as specific examples.

First, a **normed vector space** consists of an underlying vector space V over a field of scalars (the real or complex numbers), together with a **norm** $\|\cdot\|: V \to \mathbb{R}^+$ that satisfies:

- $\|v\| = 0$ if and only if $v = 0$.
- $\|\alpha v\| = |\alpha|\,\|v\|$, whenever α is a scalar and $v \in V$.
- $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.

The space V is said to be **complete** if whenever $\{v_n\}$ is a Cauchy sequence in V, that is, $\|v_n - v_m\| \to 0$ as $n, m \to \infty$, then there exists a $v \in V$ such that $\|v_n - v\| \to 0$ as $n \to \infty$.

A complete normed vector space is called a **Banach space**. Here again, we stress the importance of the fact that Cauchy sequences converge to a limit in the space itself, hence the space is "closed" under limiting operations.

3.1 Examples

The real numbers $\mathbb{R}$ with the usual absolute value form an initial example of a Banach space. Other easy examples are $\mathbb{R}^d$, with the Euclidean norm, and more generally a Hilbert space with its norm given in terms of its inner product.

Several further relevant examples are as follows:

EXAMPLE 1. The family of L^p spaces with $1 \leq p \leq \infty$ which we have just introduced are also important examples of Banach spaces (Theorem 1.3 and Theorem 2.1). Incidentally, L^2 is the only Hilbert space in the family L^p, where $1 \leq p \leq \infty$ (Exercise 25) and this in part accounts for the special flavor of the analysis carried out in L^2 as opposed to L^1 or more generally L^p for $p \neq 2$.

Finally, observe that since the triangle inequality fails in general when $0 < p < 1$, $\|\cdot\|_{L^p}$ is not a norm on L^p for this range of p, hence it is not a Banach space.

EXAMPLE 2. Another example of a Banach space is $C([0, 1])$, or more generally $C(X)$ with X a compact set in a metric space, as will be defined in Section 7. By definition, $C(X)$ is the vector space of continuous functions on X equipped with the sup-norm $\|f\| = \sup_{x\in X} |f(x)|$. Completeness is guaranteed by the fact that the uniform limit of a sequence of continuous functions is also continuous.

EXAMPLE 3. Two further examples are important in various applications. The first is the space $\Lambda^\alpha(\mathbb{R})$ of all bounded functions on $\mathbb{R}$ which satisfy a **Hölder (or Lipschitz) condition of exponent** α with $0 < \alpha \leq 1$, that is,

$$\sup_{t_1 \neq t_2} \frac{|f(t_1) - f(t_2)|}{|t_1 - t_2|^\alpha} < \infty.$$

Observe that f is then necessarily continuous; also the only interesting case is when $\alpha \leq 1$, since a function which satisfies a Hölder condition of exponent α with $\alpha > 1$ is constant.[2]

More generally, this space can be defined on $\mathbb{R}^d$; it consists of continuous functions f equipped with the norm

$$\|f\|_{\Lambda^\alpha(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d} |f(x)| + \sup_{x\neq y} \frac{|f(x) - f(y)|}{|x-y|^\alpha}.$$

With this norm, $\Lambda^\alpha(\mathbb{R}^d)$ is a Banach space (see also Exercise 29).

EXAMPLE 4. A function $f \in L^p(\mathbb{R}^d)$ is said to have **weak derivatives** in L^p up to order k, if for every multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $|\alpha| = \alpha_1 + \cdots + \alpha_d \leq k$, there is a $g_\alpha \in L^p$ with

$$\int_{\mathbb{R}^d} g_\alpha(x)\varphi(x)\,dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(x)\partial_x^\alpha \varphi(x)\,dx \tag{4}$$

for all smooth functions φ that have compact support in $\mathbb{R}^d$. Here, we use the multi-index notation

$$\partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d}.$$

Clearly, the functions g_α (when they exist) are unique, and we also write $\partial_x^\alpha f = g_\alpha$. This definition arises from the relationship (4) which holds whenever f is itself smooth, and g equals the usual derivative $\partial_x^\alpha f$, as follows from an integration by parts (see also Section 3.1, Chapter 5 in Book III).

[2]We have already encountered this space in Book I, Chapter 2 and Book III, Chapter 7.

The space $L^p_k(\mathbb{R}^d)$ is the subspace of $L^p(\mathbb{R}^d)$ of all functions that have weak derivatives up to order k. (The concept of weak derivatives will reappear in Chapter 3 in the setting of derivatives in the sense of distributions.) This space is usually referred to as a **Sobolev space**. A norm that turns $L^p_k(\mathbb{R}^d)$ into a Banach space is

$$\|f\|_{L^p_k(\mathbb{R}^d)} = \sum_{|\alpha|\leq k} \|\partial_x^\alpha f\|_{L^p(\mathbb{R}^d)} .$$

EXAMPLE 5. In the case $p = 2$, we note in the above example that an L^2 function f belongs to $L^2_k(\mathbb{R}^d)$ if and only if $(1+|\xi|^2)^{k/2}\hat{f}(\xi)$ belongs to L^2, and that $\|(1+|\xi|^2)^{k/2}\hat{f}(\xi)\|_{L^2}$ is a Hilbert space norm equivalent to $\|f\|_{L^2_k(\mathbb{R}^d)}$.

Therefore, if k is any positive number, it is natural to define L^2_k as those functions f in L^2 for which $(1+|\xi|^2)^{k/2}\hat{f}(\xi)$ belongs to L^2, and we can equip L^2_k with the norm $\|f\|_{L^2_k(\mathbb{R}^d)} = \|(1+|\xi|^2)^{k/2}\hat{f}(\xi)\|_{L^2}$.

3.2 Linear functionals and the dual of a Banach space

For the sake of simplicity, we restrict ourselves in this and the following two sections to Banach spaces over $\mathbb{R}$; the reader will find in Section 6 the slight modifications necessary to extend the results to Banach spaces over $\mathbb{C}$.

Suppose that $\mathcal{B}$ is a Banach space over $\mathbb{R}$ equipped with a norm $\|\cdot\|$. A **linear functional** is a linear mapping ℓ from $\mathcal{B}$ to $\mathbb{R}$, that is, $\ell : \mathcal{B} \to \mathbb{R}$, which satisfies

$$\ell(\alpha f + \beta g) = \alpha\ell(f) + \beta\ell(g), \qquad \text{for all } \alpha, \beta \in \mathbb{R}, \text{ and } f, g \in \mathcal{B}.$$

A linear functional ℓ is **continuous** if given $\epsilon > 0$ there exists $\delta > 0$ so that $|\ell(f) - \ell(g)| \leq \epsilon$ whenever $\|f - g\| \leq \delta$. Also we say that a linear functional is **bounded** if there is $M > 0$ with $|\ell(f)| \leq M\|f\|$ for all $f \in \mathcal{B}$. The linearity of ℓ shows that these two notions are in fact equivalent.

Proposition 3.1 *A linear functional on a Banach space is continuous, if and only if it is bounded.*

Proof. The key is to observe that ℓ is continuous if and only if ℓ is continuous at the origin.

Indeed, if ℓ is continuous, we choose $\epsilon = 1$ and $g = 0$ in the above definition so that $|\ell(f)| \leq 1$ whenever $\|f\| \leq \delta$, for some $\delta > 0$. Hence, given any non-zero h, an element of $\mathcal{B}$, we see that $\delta h/\|h\|$ has norm equal to δ, and hence $|\ell(\delta h/\|h\|)| \leq 1$. Thus $|\ell(h)| \leq M\|h\|$ with $M = 1/\delta$.

Conversely, if ℓ is bounded it is clearly continuous at the origin, hence continuous.

The significance of continuous linear functionals in terms of closed hyperplanes in $\mathcal{B}$ is a noteworthy geometric point to which we return later on. Now we take up analytic aspects of linear functionals.

The set of all continuous linear functionals over $\mathcal{B}$ is a vector space since we may add linear functionals and multiply them by scalars:

$$(\ell_1 + \ell_2)(f) = \ell_1(f) + \ell_2(f) \quad \text{and} \quad (\alpha\ell)(f) = \alpha\ell(f).$$

This vector space may be equipped with a norm as follows. The **norm** $\|\ell\|$ of a continuous linear functional ℓ is the infimum of all values M for which $|\ell(f)| \leq M\|f\|$ for all $f \in \mathcal{B}$. From this definition and the linearity of ℓ it is clear that

$$\|\ell\| = \sup_{\|f\|\leq 1} |\ell(f)| = \sup_{\|f\|=1} |\ell(f)| = \sup_{f=0} \frac{|\ell(f)|}{\|f\|}.$$

The vector space of all continuous linear functionals on $\mathcal{B}$ equipped with $\|\cdot\|$ is called the **dual space** of $\mathcal{B}$, and is denoted by $\mathcal{B}^*$.

Theorem 3.2 *The vector space $\mathcal{B}^*$ is a Banach space.*

Proof. It is clear that $\|\cdot\|$ defines a norm, so we only check that $\mathcal{B}^*$ is complete. Suppose that $\{\ell_n\}$ is a Cauchy sequence in $\mathcal{B}^*$. Then, for each $f \in \mathcal{B}$, the sequence $\{\ell_n(f)\}$ is Cauchy, hence converges to a limit, which we denote by $\ell(f)$. Clearly, the mapping $\ell : f \mapsto \ell(f)$ is linear. If M is so that $\|\ell_n\| \leq M$ for all n, we see that

$$|\ell(f)| \leq |(\ell - \ell_n)(f)| + |\ell_n(f)| \leq |(\ell - \ell_n)(f)| + M\|f\|,$$

so that in the limit as $n \to \infty$, we find $|\ell(f)| \leq M\|f\|$ for all $f \in \mathcal{B}$. Thus ℓ is bounded. Finally, we must show that ℓ_n converges to ℓ in $\mathcal{B}^*$. Given $\epsilon > 0$ choose N so that $\|\ell_n - \ell_m\| < \epsilon/2$ for all $n, m > N$. Then, if $n > N$, we see that for all $m > N$ and any f

$$|(\ell - \ell_n)(f)| \leq |(\ell - \ell_m)(f)| + |(\ell_m - \ell_n)(f)| \leq |(\ell - \ell_m)(f)| + \frac{\epsilon}{2}\|f\|.$$

We can also choose m so large (and dependent on f) so that we also have $|(\ell - \ell_m)(f)| \leq \epsilon\|f\|/2$. In the end, we find that for $n > N$,

$$|(\ell - \ell_n)(f)| \leq \epsilon\|f\|.$$

This proves that $\|\ell - \ell_n\| \to 0$, as desired.

In general, given a Banach space $\mathcal{B}$, it is interesting and very useful to be able to describe its dual $\mathcal{B}^*$. This problem has an essentially complete answer in the case of the L^p spaces introduced before.

4 The dual space of L^p when $1 \le p < \infty$

Suppose that $1 \le p \le \infty$ and q is the conjugate exponent of p, that is, $1/p + 1/q = 1$. The key observation to make is the following: Hölder's inequality shows that every function $g \in L^q$ gives rise to a bounded linear functional on L^p by

$$\ell(f) = \int_X f(x)g(x)\,d\mu(x), \tag{5}$$

and that $\|\ell\| \le \|g\|_{L^q}$. Therefore, if we associate g to ℓ above, then we find that $L^q \subset (L^p)^*$ when $1 \le p \le \infty$. The main result in this section is to prove that when $1 \le p < \infty$, every linear functional on L^p is of the form (5) for some $g \in L^q$. This implies that $(L^p)^* = L^q$ whenever $1 \le p < \infty$. We remark that this result is in general not true when $p = \infty$; the dual of L^∞ contains L^1, but it is larger. (See the end of Section 5.3 below.)

Theorem 4.1 *Suppose* $1 \le p < \infty$, *and* $1/p + 1/q = 1$. *Then, with* $\mathcal{B} = L^p$ *we have*

$$\mathcal{B}^* = L^q,$$

in the following sense: For every bounded linear functional ℓ *on* L^p *there is a unique* $g \in L^q$ *so that*

$$\ell(f) = \int_X f(x)g(x)\,d\mu(x), \qquad \textit{for all } f \in L^p.$$

Moreover, $\|\ell\|_{\mathcal{B}^*} = \|g\|_{L^q}$.

This theorem justifies the terminology whereby q is usually called the dual exponent of p.

The proof of the theorem is based on two ideas. The first, as already seen, is Hölder's inequality; to which a converse is also needed. The second is the fact that a linear functional ℓ on L^p, $1 \le p < \infty$, leads naturally to a (signed) measure ν. Because of the continuity of ℓ the measure ν is absolutely continuous with respect to the underlying measure μ, and our desired function g is then the density function of ν in terms of μ.

We begin with:

Lemma 4.2 *Suppose $1 \leq p, q \leq \infty$, are conjugate exponents.*

(i) *If $g \in L^q$, then* $\|g\|_{L^q} = \sup_{\|f\|_{L^p} \leq 1} \left| \int fg \right|$.

(ii) *Suppose g is integrable on all sets of finite measure, and*

$$\sup_{\substack{\|f\|_{L^p} \leq 1 \\ f \text{ simple}}} \left| \int fg \right| = M < \infty.$$

Then $g \in L^q$, and $\|g\|_{L^q} = M$.

For the proof of the lemma, we recall the **signum** of a real number defined by

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Proof. We start with (i). If $g = 0$, there is nothing to prove, so we may assume that g is not 0 a.e., and hence $\|g\|_{L^q} \neq 0$. By Hölder's inequality, we have that

$$\|g\|_{L^q} \geq \sup_{\|f\|_{L^p} \leq 1} \left| \int fg \right|.$$

To prove the reverse inequality we consider several cases.

- First, if $q = 1$ and $p = \infty$, we may take $f(x) = \operatorname{sign} g(x)$. Then, we have $\|f\|_{L^\infty} = 1$, and clearly, $\int fg = \|g\|_{L^1}$.

- If $1 < p, q < \infty$, then we set $f(x) = |g(x)|^{q-1} \operatorname{sign} g(x) / \|g\|_{L^q}^{q-1}$. We observe that $\|f\|_{L^p}^p = \int |g(x)|^{p(q-1)}\, d\mu / \|g\|_{L^q}^{p(q-1)} = 1$ since $p(q-1) = q$, and that $\int fg = \|g\|_{L^q}$.

- Finally, if $q = \infty$ and $p = 1$, let $\epsilon > 0$, and E a set of finite positive measure, where $|g(x)| \geq \|g\|_{L^\infty} - \epsilon$. (Such a set exists by the definition of $\|g\|_{L^\infty}$ and the fact that the measure μ is σ-finite.) Then, if we take $f(x) = \chi_E(x) \operatorname{sign} g(x)/\mu(E)$, where χ_E denotes the characteristic function of the set E, we see that $\|f\|_{L^1} = 1$, and also

$$\left| \int fg \right| = \frac{1}{\mu(E)} \int_E |g| \geq \|g\|_\infty - \epsilon.$$

This completes the proof of part (i).

To prove (ii) we recall[3] that we can find a sequence $\{g_n\}$ of simple functions so that $|g_n(x)| \leq |g(x)|$ while $g_n(x) \to g(x)$ for each x. When $p > 1$ (so $q < \infty$), we take $f_n(x) = |g_n(x)|^{q-1} \operatorname{sign} g(x)/\|g_n\|_{L^q}^{q-1}$. As before, $\|f_n\|_{L^p} = 1$. However

$$\int f_n g = \frac{\int |g_n(x)|^q}{\|g_n\|_{L^q}^{q-1}} = \|g_n\|_{L^q},$$

and this does not exceed M. By Fatou's lemma it follows that $\int |g|^q \leq M^q$, so $g \in L^q$ with $\|g\|_{L^q} \leq M$. The direction $\|g\|_{L^q} \geq M$ is of course implied by Hölder's inequality.

When $p = 1$ the argument is parallel with the above but simpler. Here we take $f_n(x) = (\operatorname{sign} g(x))\chi_{E_n}(x)$, where E_n is an increasing sequence of sets of finite measure whose union is X. The details may be left to the reader.

With the lemma established we turn to the proof of the theorem. It is simpler to consider first the case when the underlying space has finite measure. In this case, with ℓ the given functional on L^p, we can then define a set function ν by

$$\nu(E) = \ell(\chi_E),$$

where E is any measurable set. This definition makes sense because χ_E is now automatically in L^p since the space has finite measure. We observe that

$$|\nu(E)| \leq c(\mu(E))^{1/p}, \tag{6}$$

where c is the norm of the linear functional, taking into account the fact that $\|\chi_E\|_{L^p} = (\mu(E))^{1/p}$.

Now the linearity of ℓ clearly implies that ν is finitely-additive. Moreover, if $\{E_n\}$ is a countable collection of disjoint measurable sets, and we put $E = \bigcup_{n=1}^{\infty} E_n$, $E_N^* = \bigcup_{n=N+1}^{\infty} E_n$, then obviously

$$\chi_E = \chi_{E_N^*} + \sum_{n=1}^{N} \chi_{E_n}.$$

Thus $\nu(E) = \nu(E_N^*) + \sum_{n=1}^{N} \nu(E_n)$. However $\nu(E_N^*) \to 0$, as $N \to \infty$, because of (6) and the assumption $p < \infty$. This shows that ν is countably

[3] See for instance Section 2 in Chapter 6 of Book III.

additive and, moreover, (6) also shows us that ν is absolutely continuous with respect to μ.

We can now invoke the key result about absolutely continuous measures, the Lebesgue-Radon-Nykodim theorem. (See for example Theorem 4.3, Chapter 6 in Book III.) It guarantees the existence of an integrable function g so that $\nu(E) = \int_E g\,d\mu$ for every measurable set E. Thus we have $\ell(\chi_E) = \int \chi_E g\,d\mu$. The representation $\ell(f) = \int fg\,d\mu$ then extends immediately to simple functions f, and by a passage to the limit, to all $f \in L^p$ since the simple functions are dense in L^p, $1 \leq p < \infty$. (See Exercise 6.) Also by Lemma 4.2, we see that $\|g\|_{L^q} = \|\ell\|$.

To pass from the situation where the measure of X is finite to the general case, we use an increasing sequence $\{E_n\}$ of sets of finite measure that exhaust X, that is, $X = \bigcup_{n=1}^\infty E_n$. According to what we have just proved, for each n there is an integrable function g_n on E_n (which we can set to be zero in E_n^c) so that

$$\ell(f) = \int fg_n\,d\mu \tag{7}$$

whenever f is supported in E_n and $f \in L^p$. Moreover by conclusion (ii) of the lemma $\|g_n\|_{L^q} \leq \|\ell\|$.

Now it is easy to see because of (7) that $g_n = g_m$ a.e. on E_m, whenever $n \geq m$. Thus $\lim_{n\to\infty} g_n(x) = g(x)$ exists for almost every x, and by Fatou's lemma, $\|g\|_{L^q} \leq \|\ell\|$. As a result we have that $\ell(f) = \int fg\,d\mu$ for each $f \in L^p$ supported in E_n, and then by a simple limiting argument, for all $f \in L^p$. The fact that $\|\ell\| \leq \|g\|_{L^q}$, is already contained in Hölder's inequality, and therefore the proof of the theorem is complete.

5 More about linear functionals

First we turn to the study of certain geometric aspects of linear functionals in terms of the hyperplanes that they define. This will also involve understanding some elementary ideas about convexity.

5.1 Separation of convex sets

Although our ultimate focus will be on Banach spaces, we begin by considering an arbitrary vector space V over the reals. In this general setting we can define the following notions.

First, a **proper hyperplane** is a linear subspace of V that arises as the zero set of a (non-zero) linear functional on V. Alternatively, it is a linear subspace of V so that it, together with any vector not in V, spans V. Related to this notion is that of an **affine hyperplane** (which for brevity we will always refer to as a **hyperplane**) defined to be a translate of a proper hyperplane by a vector in V. To put it another way: H is a hyperplane if there is a non-zero linear functional ℓ, and a real number a, so that

$$H = \{v \in V : \quad \ell(v) = a\}.$$

Another relevant notion is that of a convex set. The subset $K \subset V$ is said to be **convex** if whenever v_0 and v_1 are both in K then the straight-line segment joining them

$$v(t) = (1-t)v_0 + tv_1, \qquad 0 \leq t \leq 1 \tag{8}$$

also lies entirely in K.

A key heuristic idea underlying our considerations can be enunciated as the following general principle:

> *If K is a convex set and $v_0 \notin K$, then K and v_0 can be separated by a hyperplane.*

This principle is illustrated in Figure 1.

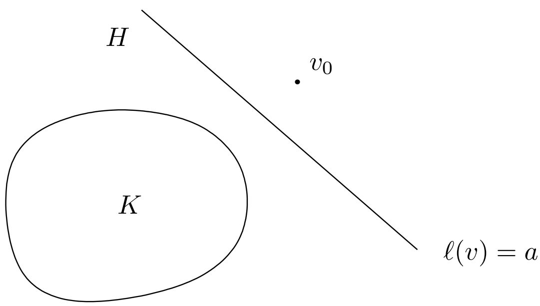

Figure 1. Separation of a convex set and a point by a hyperplane

The sense in which this is meant is that there is a non-zero linear functional ℓ and a real number a, so that

$$\ell(v_0) \geq a, \qquad \text{while} \qquad \ell(v) < a \text{ if } v \in K.$$

To give an idea of what is behind this principle we show why it holds in a nice special case. (See also Section 5.2.)

Proposition 5.1 *The assertion above is valid if $V = \mathbb{R}^d$ and K is convex and open.*

Proof. Since we may assume that K is non-empty, we can also suppose that (after a possible translation of K and v_0) we have $0 \in K$. The key construct used will be that of the Minkowski **gauge function** p associated to K, which measures (the inverse of) how far we need to go, starting from 0 in the direction of a vector v, to reach the exterior of K. The precise definition of p is as follows:

$$p(v) = \inf_{r>0}\{r : \quad v/r \in K\}.$$

Observe that since we have assumed that the origin is an interior point of K, for each $v \in \mathbb{R}^d$ there is an $r > 0$, so that $v/r \in K$. Hence $p(v)$ is well-defined.

Figure 2 below gives an example of a gauge function in the special case where $V = \mathbb{R}$ and $K = (a, b)$, an open interval that contains the origin.

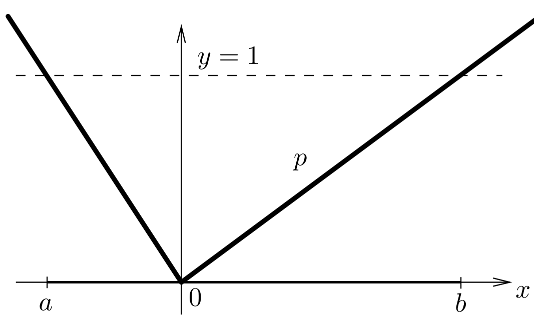

Figure 2. The gauge function of the interval (a, b) in $\mathbb{R}$

We note, for example, that if V is normed and K is the unit ball $\{\|v\| < 1\}$, then $p(v) = \|v\|$.

In general, the non-negative function p completely characterizes K in that

$$p(v) < 1 \quad \text{if and only if } v \in K. \tag{9}$$

Moreover p has an important sub-linear property:

$$\begin{cases} p(av) = ap(v), & \text{if } a \geq 0, \text{ and } v \in V. \\ p(v_1 + v_2) \leq p(v_1) + p(v_2), & \text{if } v_1 \text{ and } v_2 \in V. \end{cases} \tag{10}$$

In fact, if $v \in K$ then $v/(1-\epsilon) \in K$ for some $\epsilon > 0$, since K is open, which gives that $p(v) < 1$. Conversely if $p(v) < 1$, then $v = (1-\epsilon)v'$, for some $0 < \epsilon < 1$, and $v' \in K$. Then since $v = (1-\epsilon)v' + \epsilon \cdot 0$ this shows $v \in K$, because $0 \in K$ and K is convex.

To verify (10) we merely note that $(v_1 + v_2)/(r_1 + r_2)$ belongs to K, if both v_1/r_1 and v_2/r_2 belong to K, in view of property (8) defining the convexity of K with $t = r_2/(r_1 + r_2)$ and $1 - t = r_1/(r_1 + r_2)$.

Now our proposition will be proved once we find a linear functional ℓ, so that

$$\ell(v_0) = 1, \quad \text{and} \quad \ell(v) \le p(v), \quad v \in \mathbb{R}^d. \tag{11}$$

This is because $\ell(v) < 1$, for all $v \in K$ by (9). We shall construct ℓ in a step-by-step manner.

First, such an ℓ is already determined in the one-dimensional subspace V_0 spanned by v_0, $V_0 = \{\mathbb{R}v_0\}$, since $\ell(bv_0) = b\ell(v_0) = b$, when $b \in \mathbb{R}$, and this is consistent with (11). Indeed, if $b \ge 0$ then $p(bv_0) = bp(v_0) \ge b\ell(v_0) = \ell(bv_0)$ by (10) and (9), while (11) is immediate when $b < 0$.

The next step is to choose any vector v_1 linearly independent from v_0 and extend ℓ to the subspace V_1 spanned by v_0 and v_1. Thus we can make a choice for the value of ℓ on v_1, $\ell(v_1)$, so as to satisfy (11) if

$$a\ell(v_1) + b = \ell(av_1 + bv_0) \le p(av_1 + bv_0), \quad \text{for all } a,\, b \in \mathbb{R}.$$

Setting $a = 1$ and $bv_0 = w$ yields

$$\ell(v_1) + \ell(w) \le p(v_1 + w) \quad \text{for all } w \in V_0,$$

while setting $a = -1$ implies

$$-\ell(v_1) + \ell(w') \le p(-v_1 + w'), \quad \text{for all } w' \in V_0.$$

Altogether then it is required that for all w, $w' \in V_0$

$$-p(-v_1 + w') + \ell(w') \le \ell(v_1) \le p(v_1 + w) - \ell(w). \tag{12}$$

Notice that there is a number that lies between the two extremes of the above inequality. This is a consequence of the fact that $-p(-v_1 + w') + \ell(w')$ never exceeds $p(v_1 + w) - \ell(w)$, which itself follows from the fact that $\ell(w) + \ell(w') \le p(w + w') \le p(-v_1 + w') + p(v_1 + w)$, by (11) on V_0 and the sub-linearity of p. So a choice of $\ell(v_1)$ can be made that is consistent with (12) and this allows one to extend ℓ to V_1. In the same way we can proceed inductively to extend ℓ to all of $\mathbb{R}^d$.

The argument just given here in this special context will now be carried over in a general setting to give us an important theorem about constructing linear functionals.

5.2 The Hahn-Banach Theorem

We return to the general situation where we deal with an arbitrary vector space V over the reals. We assume that with V we are given a real-valued function p on V that satisfies the sub-linear property (10). However, as opposed to the example of the gauge function considered above, which by its nature is non-negative, here we do not assume that p has this property. In fact, certain p's which may take on negative values are needed in some of our applications later.

Theorem 5.2 *Suppose V_0 is a linear subspace of V, and that we are given a linear functional ℓ_0 on V_0 that satisfies*

$$\ell_0(v) \leq p(v), \qquad \text{for all } v \in V_0.$$

Then ℓ_0 can be extended to a linear functional ℓ on V that satisfies

$$\ell(v) \leq p(v), \qquad \text{for all } v \in V.$$

Proof. Suppose $V_0 \neq V$, and pick v_1 a vector not in V_0. We will first extend ℓ_0 to the subspace V_1 spanned by V_0 and v_1, as we did before. We can do this by defining a putative extension ℓ_1 of ℓ_0, defined on V_1 by $\ell_1(\alpha v_1 + w) = \alpha\ell_1(v_1) + \ell_0(w)$, whenever $w \in V_0$ and $\alpha \in \mathbb{R}$, if $\ell_1(v_1)$ is chosen so that

$$\ell_1(v) \leq p(v), \qquad \text{for all } v \in V_1.$$

However, exactly as above, this happens when

$$-p(-v_1 + w') + \ell_0(w') \leq \ell_1(v_1) \leq p(v_1 + w) - \ell_0(w)$$

for all w, $w' \in V_0$.

The right-hand side exceeds the left-hand side because of $\ell_0(w') + \ell_0(w) \leq p(w' + w)$ and the sub-linearity of p. Thus an appropriate choice of $\ell_1(v_1)$ is possible, giving the desired extension of ℓ_0 from V_0 to V_1.

We can think of the extension we have constructed as the key step in an inductive procedure. This induction, which in general is necessarily trans-finite, proceeds as follows. We well-order all vectors in V that do not belong to V_0, and denote this ordering by $<$. Among these vectors we call a vector v "extendable" if the linear functional ℓ_0 has an extension of the kind desired to the subspace spanned by V_0, v, and all vectors $< v$. What we want to prove is in effect that all vectors not in V_0 are extendable. Assume the contrary, then because of the well-ordering we can find the smallest v_1 that is not extendable. Now if V_0' is the space spanned by V_0 and all the vectors $< v_1$, then by assumption ℓ_0 extends to V_0'. The previous step, with V_0' in place of V_0 allows us then to extend ℓ_0 to the subspace spanned by V_0' and v_1, reaching a contradiction. This proves the theorem.

5.3 Some consequences

The Hahn-Banach theorem has several direct consequences for Banach spaces. Here $\mathcal{B}^*$ denotes the dual of the Banach space $\mathcal{B}$ as defined in Section 3.2, that is, the space of continuous linear functionals on $\mathcal{B}$.

Proposition 5.3 *Suppose f_0 is a given element of $\mathcal{B}$ with $\|f_0\| = M$. Then there exists a continuous linear functional ℓ on $\mathcal{B}$ so that $\ell(f_0) = M$ and $\|\ell\|_{\mathcal{B}^*} = 1$.*

Proof. Define ℓ_0 on the one-dimensional subspace $\{\alpha f_0\}_{\alpha\in\mathbb{R}}$ by $\ell_0(\alpha f_0) = \alpha M$, for each $\alpha \in \mathbb{R}$. Note that if we set $p(f) = \|f\|$ for every $f \in \mathcal{B}$, the function p satisfies the basic sub-linear property (10). We also observe that

$$|\ell_0(\alpha f_0)| = |\alpha| M = |\alpha| \|f_0\| = p(\alpha f_0),$$

so $\ell_0(f) \leq p(f)$ on this subspace. By the extension theorem ℓ_0 extends to an ℓ defined on $\mathcal{B}$ with $\ell(f) \leq p(f) = \|f\|$, for all $f \in \mathcal{B}$. Since this inequality also holds for $-f$ in place of f we get $|\ell(f)| \leq \|f\|$, and thus $\|\ell\|_{\mathcal{B}^*} \leq 1$. The fact that $\|\ell\|_{\mathcal{B}^*} \geq 1$ is implied by the defining property $\ell(f_0) = \|f_0\|$, thereby proving the proposition.

Another application is to the duality of linear transformations. Suppose $\mathcal{B}_1$ and $\mathcal{B}_2$ are a pair of Banach spaces, and T is a bounded linear transformation from $\mathcal{B}_1$ to $\mathcal{B}_2$. By this we mean that T maps $\mathcal{B}_1$ to $\mathcal{B}_2$; it satisfies $T(\alpha f_1 + \beta f_2) = \alpha T(f_1) + \beta T(f_2)$ whenever $f_1, f_2 \in \mathcal{B}$ and α and β are real numbers; and that it has a bound M so that $\|T(f)\|_{\mathcal{B}_2} \leq M\|f\|_{\mathcal{B}_1}$ for all $f \in \mathcal{B}_1$. The least M for which this inequality holds is called the **norm** of T and is denoted by $\|T\|$.

Often a linear transformation is initially given on a dense subspace. In this connection, the following proposition is very useful.

Proposition 5.4 *Let $\mathcal{B}_1$, $\mathcal{B}_2$ be a pair of Banach spaces and $\mathcal{S} \subset \mathcal{B}_1$ a dense linear subspace of $\mathcal{B}_1$. Suppose T_0 is a linear transformation from $\mathcal{S}$ to $\mathcal{B}_2$ that satisfies $\|T_0(f)\|_{\mathcal{B}_2} \leq M\|f\|_{\mathcal{B}_1}$ for all $f \in \mathcal{S}$. Then T_0 has a unique extension T to all of $\mathcal{B}_1$ so that $\|T(f)\|_{\mathcal{B}_2} \leq M\|f\|_{\mathcal{B}_1}$ for all $f \in \mathcal{B}_1$.*

Proof. If $f \in \mathcal{B}_1$, let $\{f_n\}$ be a sequence in $\mathcal{S}$ which converges to f. Then since $\|T_0(f_n) - T_0(f_m)\|_{\mathcal{B}_2} \leq M\|f_n - f_m\|_{\mathcal{B}_1}$ it follows that $\{T_0(f_n)\}$ is a Cauchy sequence in $\mathcal{B}_2$, and hence converges to a limit, which we define to be $T(f)$. Note that the definition of $T(f)$ is independent of the chosen sequence $\{f_n\}$, and that the resulting transformation T has all the required properties.

We now discuss duality of linear transformations. Whenever we have a linear transformation T from a Banach space $\mathcal{B}_1$ to another Banach space $\mathcal{B}_2$, it induces a **dual transformation**, T^* of $\mathcal{B}_2^*$ to $\mathcal{B}_1^*$, that can be defined as follows.

Suppose $\ell_2 \in \mathcal{B}_2^*$, (a continuous linear functional on $\mathcal{B}_2$), then $\ell_1 = T^*(\ell_2) \in \mathcal{B}_1^*$, is defined by $\ell_1(f_1) = \ell_2(T(f_1))$, whenever $f_1 \in \mathcal{B}_1$. More succinctly

$$T^*(\ell_2)(f_1) = \ell_2(T(f_1)). \tag{13}$$

Theorem 5.5 *The operator T^* defined by* (13) *is a bounded linear transformation from $\mathcal{B}_2^*$ to $\mathcal{B}_1^*$. Its norm $\|T^*\|$ satisfies $\|T\| = \|T^*\|$.*

Proof. First, if $\|f_1\|_{\mathcal{B}_1} \leq 1$, we have that

$$|\ell_1(f_1)| = |\ell_2(T(f_1))| \leq \|\ell_2\| \, \|T(f_1)\|_{\mathcal{B}_2} \leq \|\ell_2\| \, \|T\|.$$

Thus taking the supremum over all $f_1 \in \mathcal{B}_1$ with $\|f_1\|_{\mathcal{B}_1} \leq 1$, we see that the mapping $\ell_2 \mapsto T^*(\ell_2) = \ell_1$ has norm $\leq \|T\|$.

To prove the reverse inequality we can find for any $\epsilon > 0$ an $f_1 \in \mathcal{B}_1$ with $\|f_1\|_{\mathcal{B}_1} = 1$ and $\|T(f_1)\|_{\mathcal{B}_2} \geq \|T\| - \epsilon$. Next, with $f_2 = T(f_1) \in \mathcal{B}_2$, by Proposition 5.3 (with $\mathcal{B} = \mathcal{B}_2$) there is an ℓ_2 in $\mathcal{B}_2^*$ so that $\|\ell_2\|_{\mathcal{B}_2^*} = 1$ but $\ell_2(f_2) \geq \|T\| - \epsilon$. Thus by (13) one has $T^*(\ell_2)(f_1) \geq \|T\| - \epsilon$, and since $\|f_1\|_{\mathcal{B}_1} = 1$, we conclude $\|T^*(\ell_2)\|_{\mathcal{B}_1^*} \geq \|T\| - \epsilon$. This gives $\|T^*\| \geq \|T\| - \epsilon$ for any $\epsilon > 0$, which proves the theorem.

A further quick application of the Hahn-Banach theorem is the observation that in general L^1 is *not* the dual of L^∞ (as opposed to the case $1 \leq p < \infty$ considered in Theorem 4.1).

Let us first recall that whenever $g \in L^1$, the linear functional $f \mapsto \ell(f)$ given by

$$\ell(f) = \int fg \, d\mu \tag{14}$$

is bounded on L^∞, and its norm $\|\ell\|_{(L^\infty)^*}$ is $\|g\|_{L^1}$. In this way L^1 can be viewed as a subspace of $(L^\infty)^*$, with the L^1 norm of g being identical with its norm as a linear functional. One can, however, produce a continuous linear functional of L^∞ not of this form. For simplicity we do this when the underlying space is $\mathbb{R}$ with Lebesgue measure.

We let $\mathcal{C}$ denote the subspace of $L^\infty(\mathbb{R})$ consisting of continuous bounded functions on $\mathbb{R}$. Define the linear function ℓ_0 on $\mathcal{C}$ (the "Dirac delta") by

$$\ell_0(f) = f(0), \qquad f \in \mathcal{C}.$$

Clearly $|\ell_0(f)| \leq \|f\|_{L^\infty}$, $f \in \mathcal{C}$. Thus by the extension theorem, with $p(f) = \|f\|_{L^\infty}$, we see that there is a linear functional ℓ on L^∞, extending ℓ_0, that satisfies $|\ell(f)| \leq \|f\|_{L^\infty}$, for all $f \in L^\infty$.

Suppose for a moment that ℓ were of the form (14) for some $g \in L^1$. Since $\ell(f) = \ell_0(f) = 0$ whenever f is a continuous trapezoidal function that excludes the origin, we would have $\int fg \, dx = 0$ for such functions f; by a simple limiting argument this gives $\int_I g \, dx = 0$ for all intervals excluding the origin, and from there for all intervals I. Hence the indefinite integrals $G(y) = \int_0^y g(x) \, dx$ vanish, and therefore $G' = g = 0$ by the differentiation theorem.[4] This gives a contradiction, hence the linear functional ℓ is not representable as (14).

5.4 The problem of measure

We now consider an application of the Hahn-Banach theorem of a different kind. We present a rather stunning assertion, answering a basic question of the "problem of measure." The result states that there is a finitely-additive[5] measure defined on *all* subsets of $\mathbb{R}^d$ that agrees with Lebesgue measure on the measurable sets, and is translation invariant. We formulate the theorem in one dimension.

Theorem 5.6 *There is an extended-valued non-negative function $\hat{m}$, defined on all subsets of $\mathbb{R}$ with the following properties:*

(i) $\hat{m}(E_1 \cup E_2) = \hat{m}(E_1) + \hat{m}(E_2)$ *whenever E_1 and E_2 are disjoint subsets of $\mathbb{R}$.*

[4]See for instance Theorem 3.11, in Chapter 3 of Book III.
[5]The qualifier "finitely-additive" is crucial.

(ii) $\hat{m}(E) = m(E)$ *if* E *is a measurable set and* m *denotes the Lebesgue measure.*

(iii) $\hat{m}(E + h) = \hat{m}(E)$ *for every set* E *and real number* h.

From (i) we see that $\hat{m}$ is finitely additive; however it cannot be countably additive as the proof of the existence of non-measurable sets shows. (See Section 3, Chapter 1 in Book III.)

This theorem is a consequence of another result of this kind, dealing with an extension of the Lebesgue integral. Here the setting is the circle $\mathbb{R}/\mathbb{Z}$, instead of $\mathbb{R}$, with the former realized as $(0, 1]$. Thus functions on $\mathbb{R}/\mathbb{Z}$ can be thought of as functions on $(0, 1]$, extended to $\mathbb{R}$ by periodicity with period 1. In the same way, translations on $\mathbb{R}$ induce corresponding translations on $\mathbb{R}/\mathbb{Z}$. The assertion now is the existence of a generalized integral (the "Banach integral") defined on all bounded functions on the circle.

Theorem 5.7 *There is a linear functional* $f \mapsto I(f)$ *defined on all bounded functions* f *on* $\mathbb{R}/\mathbb{Z}$ *so that:*

(a) $I(f) \geq 0$, *if* $f(x) \geq 0$ *for all* x.

(b) $I(\alpha f_1 + \beta f_2) = \alpha I(f_1) + \beta I(f_2)$ *for all* α *and* β *real.*

(c) $I(f) = \int_0^1 f(x)\, dx$, *whenever* f *is measurable.*

(d) $I(f_h) = I(f)$, *for all* $h \in \mathbb{R}$ *where* $f_h(x) = f(x - h)$.

The right-hand side of (c) denotes the usual Lebesgue integral.

Proof. The idea is to consider the vector space V of all (real-valued) bounded functions on $\mathbb{R}/\mathbb{Z}$, with V_0 the subspace of those functions that are measurable. We let I_0 denote the linear functional given by the Lebesgue integral, $I_0(f) = \int_0^1 f(x)\, dx$ for $f \in V_0$. The key is to find the appropriate sub-linear p defined on V so that

$$I_0(f) \leq p(f), \qquad \text{for all } f \in V_0.$$

Banach's ingenious definition of p is as follows: We let $A = \{a_1, \ldots, a_N\}$ denote an arbitrary collection of N real numbers, with $\#(A) = N$ denoting its cardinality. Given A, we define $M_A(f)$ to be the real number

$$M_A(f) = \sup_{x \in \mathbb{R}} \frac{1}{N} \sum_{j=1}^{N} f(x + a_j) \quad ,$$

and set

$$p(f) = \inf_A \{M_A(f)\},$$

where the infimum is taken over all finite collections A.

It is clear that $p(f)$ is well-defined, since f is assumed to be bounded; also $p(cf) = cp(f)$ if $c \geq 0$. To prove $p(f_1 + f_2) \leq p(f_1) + p(f_2)$, we find for each ϵ, finite collections A and B so that

$$M_A(f_1) \leq p(f_1) + \epsilon \quad \text{and} \quad M_B(f_2) \leq p(f_2) + \epsilon.$$

Let C be the collection $\{a_i + b_j\}_{1 \leq i \leq N_1,\ 1 \leq j \leq N_2}$ where $N_1 = \#(A)$, and $N_2 = \#(B)$. Now it is easy to see that

$$M_C(f_1 + f_2) \leq M_C(f_1) + M_C(f_2).$$

Next, we note as a general matter that $M_A(f)$ is the same as $M_{A'}(f')$ where $f' = f_h$ is a translate of f and $A' = A - h$. Also the averages corresponding to C arise as averages of translates of the averages corresponding to A and B, so it is easy to verify that

$$M_C(f_1) \leq M_A(f_1) \quad \text{and also} \quad M_C(f_2) \leq M_B(f_2).$$

Thus

$$p(f_1 + f_2) \leq M_C(f_1 + f_2) \leq M_A(f_1) + M_B(f_2) \leq p(f_1) + p(f_2) + 2\epsilon.$$

Letting $\epsilon \to 0$ proves the sub-linearity of p.

Next if f is Lebesgue measurable (and hence integrable since it is bounded), then for each A

$$I_0(f) = \frac{1}{N} \int_0^1 \sum_{j=1}^{N} f(x + a_j) \quad dx \leq \int_0^1 M_A(f)\, dx = M_A(f),$$

and hence $I_0(f) \leq p(f)$. Let therefore I be the linear functional extending I_0 from V_0 to V, whose existence is guaranteed by Theorem 5.2. It is obvious from its definition that $p(f) \leq 0$ if $f \leq 0$. From this it follows that $I(f) \leq 0$ when $f \leq 0$, and replacing f by $-f$ we see that conclusion (a) holds.

Next we observe that for each real h

$$p(f - f_h) \leq 0. \tag{15}$$

In fact, for h fixed and N given, define the set A_N to be $\{h, 2h, 3h, \ldots, Nh\}$. Then the sum that enters in the definition of $M_{A_N}(f - f_h)$ is

$$\frac{1}{N}\sum_{j=1}^{N}\left(f(x+jh) - f(x+(j-1)h)\right),$$

and thus $|M_{A_N}(f - f_h)| \leq 2M/N$, where M is an upper bound for $|f|$. Since $p(f - f_h) \leq M_{A_N}(f - f_h) \to 0$, as $N \to \infty$, we see that (15) is proved. This shows that $I(f - f_h) \leq 0$, for all f and h. However, replacing f by f_h and then h by $-h$, we see that $I(f_h - f) \leq 0$ and thus (d) is also established, finishing the proof of Theorem 5.7.

As a direct consequence we have the following.

Corollary 5.8 *There is a non-negative function $\hat{m}$ defined on all subsets of $\mathbb{R}/\mathbb{Z}$ so that:*

(i) $\hat{m}(E_1 \cup E_2) = \hat{m}(E_1) + \hat{m}(E_2)$ *for all disjoint subsets E_1 and E_2.*

(ii) $\hat{m}(E) = m(E)$ *if E is measurable.*

(iii) $\hat{m}(E + h) = \hat{m}(E)$ *for every h in $\mathbb{R}$.*

We need only take $\hat{m}(E) = I(\chi_E)$, with I as in Theorem 5.7, where χ_E denotes the characteristic function of E.

We now turn to the proof of Theorem 5.6. Let $\mathcal{I}_j$ denote the interval $(j, j+1]$, where $j \in \mathbb{Z}$. Then we have a partition $\bigcup_{j=-\infty}^{\infty} \mathcal{I}_j$ of $\mathbb{R}$ into disjoint sets.

For clarity of exposition, we temporarily relabel the measure $\hat{m}$ on $(0, 1] = \mathcal{I}_0$ given by the corollary and call it $\hat{m}_0$. So whenever $E \subset \mathcal{I}_0$ we defined $\hat{m}(E)$ to be $\hat{m}_0(E)$. More generally, if $E \subset \mathcal{I}_j$ we set $\hat{m}(E) = \hat{m}_0(E - j)$.

With these things said, for any set E define $\hat{m}(E)$ by

$$\hat{m}(E) = \sum_{j=-\infty}^{\infty} \hat{m}(E \cap \mathcal{I}_j) = \sum_{j=-\infty}^{\infty} \hat{m}_0((E \cap \mathcal{I}_j) - j). \tag{16}$$

Thus $\hat{m}(E)$ is given as an extended non-negative number. Note that if E_1 and E_2 are disjoint so are $(E_1 \cap \mathcal{I}_j) - j$ and $(E_2 \cap \mathcal{I}_j) - j$. It follows that $\hat{m}(E_1 \cup E_2) = \hat{m}(E_1) + \hat{m}(E_2)$. Moreover if E is measurable then $\hat{m}(E \cap \mathcal{I}_j) = m(E \cap \mathcal{I}_j)$ and so $\hat{m}(E) = m(E)$.

To prove $\hat{m}(E + h) = \hat{m}(E)$, consider first the case $h = k \in \mathbb{Z}$. This is an immediate consequence of the definition (16) once one observes that $((E + k) \cap \mathcal{I}_{j+k}) - (j + k) = (E \cap \mathcal{I}_j) - j$, for all j, $k \in \mathbb{Z}$.

Next suppose $0 < h < 1$. We then decompose $E \cap \mathcal{I}_j$ as $E'_j \cup E''_j$, with $E'_j = E \cap (j, j+1-h]$ and $E''_j = E \cap (j+1-h, j+1]$. The point of this decomposition is that $E'_j + h$ remains in $\mathcal{I}_j$ but $E''_j + h$ is placed in $\mathcal{I}_{j+1}$. In any case, $E = \bigcup_j E'_j \cup \bigcup_j E''_j$, and the union is disjoint.

Thus using the first additivity property proved above and then (16) we see that

$$\hat{m}(E) = \sum_{j=-\infty}^{\infty} \left(\hat{m}(E'_j) + \hat{m}(E''_j)\right).$$

Similarly

$$\hat{m}(E+h) = \sum_{j=-\infty}^{\infty} \left(\hat{m}(E'_j + h) + \hat{m}(E''_j + h)\right).$$

Now both E'_j and $E'_j + h$ are in $\mathcal{I}_j$, hence $\hat{m}(E'_j) = \hat{m}(E'_j + h)$ by the translation invariance of $\hat{m}_0$ and the definition of $\hat{m}$ on subsets of $\mathcal{I}_j$. Also E''_j is in $\mathcal{I}_j$ and $E''_j + h$ is in $\mathcal{I}_{j+1}$, and their measures agree for the same reasons. This establishes that $\hat{m}(E) = \hat{m}(E+h)$, for $0 < h < 1$. Now combining this with the translation invariance with respect to $\mathbb{Z}$ already proved, we obtain conclusion (iii) of Theorem 5.6 for all h, and hence the theorem is completely proved.

For the corresponding extension of Lebesgue measure in $\mathbb{R}^d$ and other related results, see Exercise 36 and Problems 8^* and 9^*.

6 Complex L^p and Banach spaces

We have supposed in Section 3.2 onwards that our L^p and Banach spaces are taken over the reals. However, the statements and the proofs of the corresponding theorems for those spaces taken with respect to the complex scalars are for the most part routine adaptations of the real case. There are nevertheless several instances that require further comment. First, in the argument concerning the converse of Hölder's inequality (Lemma 4.2), the definition of f should read

$$f(x) = |g(x)|^{q-1} \frac{\overline{\operatorname{sign} g(x)}}{\|g\|_{L^q}^{q-1}},$$

where now "sign" denotes the complex version of the signum function, defined by $\operatorname{sign} z = z/|z|$ if $z \neq 0$, and $\operatorname{sign} 0 = 0$. There are similar occurrences with g replaced by g_n.

Second, while the Hahn-Banach theorem is valid as stated only for real vector spaces, a version of the complex case (sufficient for the applications in Section 5.3 where $p(f) = \|f\|$) can be found in Exercise 33 below.

7 Appendix: The dual of $C(X)$

In this appendix, we describe the bounded linear functionals of the space $C(X)$ of continuous real-valued functions on X. To begin with, we assume that X is a compact metric space. Our main result then states that if $\ell \in C(X)^*$, then there exists a finite signed Borel measure μ (this measure is sometimes referred to as a Radon measure) so that

$$\ell(f) = \int_X f(x)\, d\mu(x) \qquad \text{for all } f \in C(X).$$

Before proceeding with the argument leading to this result, we collect some basic facts and definitions.

Let X be a metric space with metric d, and assume that X is compact; that is, every covering of X by open sets contains a finite sub-covering. The vector space $C(X)$ of real-valued continuous functions on X equipped with the sup-norm

$$\|f\| = \sup_{x\in X} |f(x)|, \qquad f \in C(X)$$

is a Banach space over $\mathbb{R}$. Given a continuous function f on X we define the **support** of f, denoted $\mathrm{supp}(f)$, as the closure of the set $\{x \in X : f(x) \neq 0\}$.[6]

We recall some simple facts about continuous functions and open and closed sets in X that we shall use below.

(i) **Separation.** If A and B are two disjoint closed subsets of X, then there exists a continuous function f with $f = 1$ on A, $f = 0$ on B, and $0 < f < 1$ in the complements of A and B.

Indeed, one can take for instance

$$f(x) = \frac{d(x,B)}{d(x,A) + d(x,B)},$$

where $d(x,B) = \inf_{y\in B} d(x,y)$, with a similar definition for $d(x,A)$.

(ii) **Partition of unity.** If K is a compact set which is covered by finitely many open sets $\{\mathcal{O}_k\}_{k=1}^N$, then there exist continuous functions η_k for $1 \leq k \leq N$ so that $0 \leq \eta_k \leq 1$, $\mathrm{supp}(\eta_k) \subset \mathcal{O}_k$, and $\sum_{k=1}^N \eta_k(x) = 1$ whenever $x \in K$. Moreover, $0 \leq \sum_{k=1}^N \eta_k(x) \leq 1$ for all $x \in X$.

One can argue as follows. For each $x \in K$, there exists a ball $B(x)$ centered at x and of positive radius such that $\overline{B(x)} \subset \mathcal{O}_i$ for some i. Since $\bigcup_{x\in K} B(x)$ covers K, we can select a finite subcovering, say $\bigcup_{j=1}^M B(x_j)$. For each $1 \leq k \leq N$, let U_k be the union of all open balls $B(x_j)$ so that $\overline{B(x_j)} \subset \mathcal{O}_k$; clearly $K \subset \bigcup_{k=1}^N U_k$. By (i) above, there exists a continuous function $0 \leq \varphi_k \leq 1$ so that $\varphi_k = 1$ on $\overline{U_k}$ and $\mathrm{supp}(\varphi_k) \subset \mathcal{O}_k$. If we define

$$\eta_1 = \varphi_1, \quad \eta_2 = \varphi_2(1-\varphi_1), \ \ldots\ , \quad \eta_N = \varphi_N(1-\varphi_1)\cdots(1-\varphi_{N-1})$$

[6] This is the common usage of the terminology "support." In Book III, Chapter 2, we used "support of f" to indicate the set where $f(x) \neq 0$, which is convenient when dealing with measurable functions.

then $\text{supp}(\eta_k) \subset \mathcal{O}_k$ and

$$\eta_1 + \cdots + \eta_N = 1 - (1 - \varphi_1) \cdots (1 - \varphi_N),$$

thus guaranteeing the desired properties.

Recall[7] that the **Borel σ-algebra** of X, which is denoted by $\mathcal{B}_X$, is the smallest σ-algebra of X that contains the open sets. Elements of $\mathcal{B}_X$ are called **Borel sets**, and a measure defined on $\mathcal{B}_X$ is called a **Borel measure**. If a Borel measure is finite, that is $\mu(X) < \infty$, then it satisfies the following "regularity property": for any Borel set E and any $\epsilon > 0$, there are an open set $\mathcal{O}$ and a closed set F such that $E \subset \mathcal{O}$ and $\mu(\mathcal{O} - E) < \epsilon$, while $F \subset E$ and $\mu(E - F) < \epsilon$.

In general we shall be interested in finite *signed* Borel measures on X, that is, measures which can take on negative values. If μ is such a measure, and μ^+ and μ^- denote the positive and negative variations of μ, then $\mu = \mu^+ - \mu^-$, and integration with respect to μ is defined by $\int f\, d\mu = \int f\, d\mu^+ - \int f\, d\mu^-$. Conversely, if μ_1 and μ_2 are two finite Borel measures, then $\mu = \mu_1 - \mu_2$ is a finite signed Borel measure, and $\int f\, d\mu = \int f\, d\mu_1 - \int f\, d\mu_2$.

We denote by $M(X)$ the space of finite signed Borel measures on X. Clearly, $M(X)$ is a vector space which can be equipped with the following norm

$$\|\mu\| = |\mu|(X),$$

where $|\mu|$ denotes the total variation of μ. It is a simple fact that $M(X)$ with this norm is a Banach space.

7.1 The case of positive linear functionals

We begin by considering only linear functionals $\ell : C(X) \to \mathbb{R}$ which are **positive**, that is, $\ell(f) \geq 0$ whenever $f(x) \geq 0$ for all $x \in X$. Observe that positive linear functionals are automatically bounded and that $\|\ell\| = \ell(1)$. Indeed, note that $|f(x)| \leq \|f\|$, hence $\|f\| \pm f \geq 0$, and therefore $|\ell(f)| \leq \ell(1)\|f\|$.

Our main result goes as follows.

Theorem 7.1 *Suppose X is a compact metric space and ℓ a positive linear functional on $C(X)$. Then there exists a unique finite (positive) Borel measure μ so that*

$$\ell(f) = \int_X f(x)\, d\mu(x) \qquad \text{for all } f \in C(X). \tag{17}$$

Proof. The existence of the measure μ is proved as follows. Consider the function ρ on the open subsets of X defined by

$$\rho(\mathcal{O}) = \sup \{\ell(f), \text{ where } \text{supp}(f) \subset \mathcal{O}, \text{ and } 0 \leq f \leq 1\},$$

[7]The definitions and results on measure theory needed in this section, in particular the extension of a premeasure used in the proof of Theorem 7.1, can be found in Chapter 6 of Book III.

and let the function μ_* be defined on all subsets of X by

$$\mu_*(E) = \inf\{\rho(\mathcal{O}), \quad \text{where } E \subset \mathcal{O} \text{ and } \mathcal{O} \text{ is open}\}.$$

We contend that μ_* is a metric exterior measure on X.

Indeed, we clearly must have $\mu_*(E_1) \leq \mu_*(E_2)$ whenever $E_1 \subset E_2$. Also, if $\mathcal{O}$ is open, then $\mu_*(\mathcal{O}) = \rho(\mathcal{O})$. To show that μ_* is countably sub-additive on subsets of X, we begin by proving that μ_* is in fact sub-additive on open sets $\{\mathcal{O}_k\}$, that is,

$$\mu_*\left(\bigcup_{k=1}^{\infty} \mathcal{O}_k\right) \leq \sum_{k=1}^{\infty} \mu_*(\mathcal{O}_k). \tag{18}$$

To do so, suppose $\{\mathcal{O}_k\}_{k=1}^{\infty}$ is a collection of open sets in X, and let $\mathcal{O} = \bigcup_{k=1}^{\infty} \mathcal{O}_k$. If f is any continuous function that satisfies $\text{supp}(f) \subset \mathcal{O}$ and $0 \leq f \leq 1$, then by compactness of $K = \text{supp}(f)$ we can pick a sub-cover so that (after relabeling the sets $\mathcal{O}_k$, if necessary) $K \subset \bigcup_{k=1}^{N} \mathcal{O}_k$. Let $\{\eta_k\}_{k=1}^{N}$ be a partition of unity of $\{\mathcal{O}_1, \ldots, \mathcal{O}_N\}$ (as discussed above in (ii)); this means that each η_k is continuous with $0 \leq \eta_k \leq 1$, $\text{supp}(\eta_k) \subset \mathcal{O}_k$ and $\sum_{k=1}^{N} \eta_k(x) = 1$ for all $x \in K$. Hence recalling that $\mu_* = \rho$ on open sets, we get

$$\ell(f) = \sum_{k=1}^{N} \ell(f\eta_k) \leq \sum_{k=1}^{N} \mu_*(\mathcal{O}_k) \leq \sum_{k=1}^{\infty} \mu_*(\mathcal{O}_k),$$

where the first inequality follows because $\text{supp}(f\eta_k) \subset \mathcal{O}_k$ and $0 \leq f\eta_k \leq 1$. Taking the supremum over f we find that $\mu_*\left(\bigcup_{k=1}^{\infty} \mathcal{O}_k\right) \leq \sum_{k=1}^{\infty} \mu_*(\mathcal{O}_k)$.

We now turn to the proof of the sub-additivity of μ_* on all sets. Suppose $\{E_k\}$ is a collection of subsets of X and let $\epsilon > 0$. For each k, pick an open set $\mathcal{O}_k$ so that $E_k \subset \mathcal{O}_k$ and $\mu_*(\mathcal{O}_k) \leq \mu_*(E_k) + \epsilon 2^{-k}$. Since $\mathcal{O} = \bigcup \mathcal{O}_k$ covers $\bigcup E_k$, we must have by (18) that

$$\mu_*(\bigcup E_k) \leq \mu_*(\mathcal{O}) \leq \sum_k \mu_*(\mathcal{O}_k) \leq \sum_k \mu_*(E_k) + \epsilon,$$

and consequently $\mu_*(\bigcup E_k) \leq \sum_k \mu_*(E_k)$ as desired.

The last property we must verify is that μ_* is metric, in the sense that if $d(E_1, E_2) > 0$, then $\mu_*(E_1 \cup E_2) = \mu_*(E_1) + \mu_*(E_2)$. Indeed, the separation condition implies that there exist disjoint open sets $\mathcal{O}_1$ and $\mathcal{O}_2$ so that $E_1 \subset \mathcal{O}_1$ and $E_2 \subset \mathcal{O}_2$. Therefore, if $\mathcal{O}$ is any open subset which contains $E_1 \cup E_2$, then $\mathcal{O} \supset (\mathcal{O} \cap \mathcal{O}_1) \cup (\mathcal{O} \cap \mathcal{O}_2)$, where this union is disjoint. Hence the additivity of μ_* on disjoint open sets, and its monotonicity give

$$\mu_*(\mathcal{O}) \geq \mu_*(\mathcal{O} \cap \mathcal{O}_1) + \mu_*(\mathcal{O} \cap \mathcal{O}_2) \geq \mu_*(E_1) + \mu_*(E_2),$$

since $E_1 \subset (\mathcal{O} \cap \mathcal{O}_1)$ and $E_2 \subset (\mathcal{O} \cap \mathcal{O}_2)$. So $\mu_*(E_1 \cup E_2) \geq \mu_*(E_1) + \mu_*(E_2)$, and since the reverse inequality has already been shown above, this concludes the proof that μ_* is a metric exterior measure.

By Theorems 1.1 and 1.2 in Chapter 6 of Book III, there exists a Borel measure μ on $\mathcal{B}_X$ which extends μ_*. Clearly, μ is finite with $\mu(X) = \ell(1)$.

We now prove that this measure satisfies (17). Let $f \in C(X)$. Since f can be written as the difference of two continuous non-negative functions, we can assume after rescaling, that $0 \le f(x) \le 1$ for all $x \in X$. The idea now is to slice f, that is, write $f = \sum f_n$ where each f_n is continuous and relatively small in the sup-norm. More precisely, let N be a fixed positive integer, define $\mathcal{O}_0 = X$, and for every integer $n \ge 1$, let

$$\mathcal{O}_n = \{x \in X : f(x) > (n-1)/N\}.$$

Thus $\mathcal{O}_n \supset \mathcal{O}_{n+1}$ and $\mathcal{O}_{N+1} = \emptyset$. Now if we define

$$f_n(x) = \begin{cases} 1/N & \text{if } x \in \mathcal{O}_{n+1}, \\ f(x) - (n-1)/N & \text{if } x \in \mathcal{O}_n - \mathcal{O}_{n+1}, \\ 0 & \text{if } x \in \mathcal{O}_n^c, \end{cases}$$

then the functions f_n are continuous and they "pile up" to yield f, that is, $f = \sum_{n=1}^N f_n$. Since $Nf_n = 1$ on $\mathcal{O}_{n+1}$, $\operatorname{supp}(Nf_n) \subset \overline{\mathcal{O}_n} \subset \mathcal{O}_{n-1}$, and also $0 \le Nf_n \le 1$ we have $\mu(\mathcal{O}_{n+1}) \le \ell(Nf_n) \le \mu(\mathcal{O}_{n-1})$, and therefore by linearity

$$\frac{1}{N}\sum_{n=1}^{N} \mu(\mathcal{O}_{n+1}) \le \ell(f) \le \frac{1}{N}\sum_{n=1}^{N} \mu(\mathcal{O}_{n-1}). \tag{19}$$

The properties of Nf_n also imply $\mu(\mathcal{O}_{n+1}) \le \int Nf_n \, d\mu \le \mu(\mathcal{O}_n)$, hence

$$\frac{1}{N}\sum_{n=1}^{N} \mu(\mathcal{O}_{n+1}) \le \int f \, d\mu \le \frac{1}{N}\sum_{n=1}^{N} \mu(\mathcal{O}_n). \tag{20}$$

Consequently, combining the inequalities (19) and (20) yields

$$\left| \ell(f) - \int f \, d\mu \right| \le \frac{2\mu(X)}{N}.$$

In the limit as $N \to \infty$ we conclude that $\ell(f) = \int f \, d\mu$ as desired.

Finally, we prove uniqueness. Suppose μ' is another finite positive Borel measure on X that satisfies $\ell(f) = \int f \, d\mu'$ for all $f \in C(X)$. If $\mathcal{O}$ is an open set, and $0 \le f \le 1$ with $\operatorname{supp}(f) \subset \mathcal{O}$, then

$$\ell(f) = \int f \, d\mu' = \int_{\mathcal{O}} f \, d\mu' \le \int_{\mathcal{O}} 1 \, d\mu' = \mu'(\mathcal{O}).$$

Taking the supremum over f and recalling the definition of μ yields $\mu(\mathcal{O}) \le \mu'(\mathcal{O})$. For the reverse inequality, recall the inner regularity condition satisfied by a finite Borel measure: given $\epsilon > 0$, there exists a closed set K so that $K \subset \mathcal{O}$, and $\mu'(\mathcal{O} - K) < \epsilon$. By the separation property (i) noted above applied to K and $\mathcal{O}^c$, we can pick a continuous function f so that $0 \le f \le 1$, $\mathrm{supp}(f) \subset \mathcal{O}$ and $f = 1$ on K. Then

$$\mu'(\mathcal{O}) \le \mu'(K) + \epsilon \le \int_K f\, d\mu' + \epsilon \le \ell(f) + \epsilon \le \mu(\mathcal{O}) + \epsilon.$$

Since ϵ was arbitrary, we obtain the desired inequality, and therefore $\mu(\mathcal{O}) = \mu'(\mathcal{O})$ for all open sets $\mathcal{O}$. This implies that $\mu = \mu'$ on all Borel sets, and the proof of the theorem is complete.

7.2 The main result

The main point is to write an arbitrary bounded linear functional on $C(X)$ as the difference of two positive linear functionals.

Proposition 7.2 *Suppose X is a compact metric space and let ℓ be a bounded linear functional on $C(X)$. Then there exist positive linear functionals ℓ^+ and ℓ^- so that $\ell = \ell^+ - \ell^-$. Moreover, $\|\ell\| = \ell^+(1) + \ell^-(1)$.*

Proof. For $f \in C(X)$ with $f \ge 0$, we define

$$\ell^+(f) = \sup\{\ell(\varphi) : 0 \le \varphi \le f\}.$$

Clearly, we have $0 \le \ell^+(f) \le \|\ell\|\|f\|$ and $\ell(f) \le \ell^+(f)$. If $\alpha \ge 0$ and $f \ge 0$, then $\ell^+(\alpha f) = \alpha\ell^+(f)$. Now suppose that $f, g \ge 0$. On the one hand we have $\ell^+(f) + \ell^+(g) \le \ell^+(f + g)$, because if $0 \le \varphi \le f$ and $0 \le \psi \le g$, then $0 \le \varphi + \psi \le f + g$. On the other hand, suppose $0 \le \varphi \le f + g$, and let $\varphi_1 = \min(\varphi, f)$ and $\varphi_2 = \varphi - \varphi_1$. Then $0 \le \varphi_1 \le f$ and $0 \le \varphi_2 \le g$, and $\ell(\varphi) = \ell(\varphi_1) + \ell(\varphi_2) \le \ell^+(f) + \ell^+(g)$. Taking the supremum over φ, we get $\ell^+(f + g) \le \ell^+(f) + \ell^+(g)$. We conclude from the above that $\ell^+(f + g) = \ell^+(f) + \ell^+(g)$ whenever $f, g \ge 0$.

We can now extend ℓ^+ to a positive linear functional on $C(X)$ as follows. Given an arbitrary function f in $C(X)$ we can write $f = f^+ - f^-$, where $f^+, f^- \ge 0$, and define ℓ^+ on f by $\ell^+(f) = \ell^+(f^+) - \ell^+(f^-)$. Using the linearity of ℓ^+ on non-negative functions, one checks easily that the definition of $\ell^+(f)$ is independent of the decomposition of f into the difference of two non-negative functions. From the definition we see that ℓ^+ is positive, and it is easy to check that ℓ^+ is linear on $C(X)$, and that $\|\ell^+\| \le \|\ell\|$.

Finally, we define $\ell^- = \ell^+ - \ell$, and see immediately that ℓ^- is also a positive linear functional on $C(X)$.

Now since ℓ^+ and ℓ^- are positive, we have $\|\ell^+\| = \ell^+(1)$ and $\|\ell^-\| = \ell^-(1)$, therefore $\|\ell\| \le \ell^+(1) + \ell^-(1)$. For the reverse inequality, suppose $0 \le \varphi \le 1$. Then $|2\varphi - 1| \le 1$, hence $\|\ell\| \ge \ell(2\varphi - 1)$. By linearity of ℓ, and taking the supremum over φ we obtain $\|\ell\| \ge 2\ell^+(1) - \ell(1)$. Since $\ell(1) = \ell^+(1) - \ell^-(1)$ we get $\|\ell\| \ge \ell^+(1) + \ell^-(1)$, and the proof is complete.

We are now ready to state and prove the main result.

Theorem 7.3 *Let X be a compact metric space and $C(X)$ the Banach space of continuous real-valued functions on X. Then, given any bounded linear functional ℓ*

on $C(X)$, there exists a unique finite signed Borel measure μ on X so that

$$\ell(f) = \int_X f(x)\, d\mu(x) \qquad \textit{for all } f \in C(X).$$

Moreover, $\|\ell\| = \|\mu\| = |\mu|(X)$. In other words $C(X)^*$ is isometric to $M(X)$.

Proof. By the proposition, there exist two positive linear functionals ℓ^+ and ℓ^- so that $\ell = \ell^+ - \ell^-$. Applying Theorem 7.1 to each of these positive linear functionals yields two finite Borel measures μ_1 and μ_2. If we define $\mu = \mu_1 - \mu_2$, then μ is a finite signed Borel measure and $\ell(f) = \int f\, d\mu$.

Now we have

$$|\ell(f)| \leq \int |f|\, d|\mu| \leq \|f\|\, |\mu|(X),$$

and thus $\|\ell\| \leq |\mu|(X)$. Since we also have $|\mu|(X) \leq \mu_1(X) + \mu_2(X) = \ell^+(1) + \ell^-(1) = \|\ell\|$, we conclude that $\|\ell\| = |\mu|(X)$ as desired.

To prove uniqueness, suppose $\int f\, d\mu = \int f\, d\mu'$ for some finite signed Borel measures μ and μ', and all $f \in C(X)$. Then if $\nu = \mu - \mu'$, one has $\int f d\nu = 0$, and consequently, if ν^+ and ν^- are the positive and negative variations of f, one finds that the two positive linear functionals defined on $C(X)$ by $\ell^+(f) = \int f\, d\nu^+$ and $\ell^-(f) = \int f\, d\nu^-$ are identical. By the uniqueness in Theorem 7.1, we conclude that $\nu^+ = \nu^-$, hence $\nu = 0$ and $\mu = \mu'$, as desired.

7.3 An extension

Because of its later application, it is useful to observe that Theorem 7.1 has an extension when we drop the assumption that the space X is compact. Here we define the space $C_b(X)$ of continuous bounded functions f on X, with norm $\|f\| = \sup_{x\in X} |f(x)|$.

Theorem 7.4 *Suppose X is a metric space and ℓ a positive linear functional on $C_b(X)$. For simplicity assume that ℓ is normalized so that $\ell(1) = 1$. Assume also that for each $\epsilon > 0$, there is a compact set $K_\epsilon \subset X$ so that*

$$|\ell(f)| \leq \sup_{x\in K_\epsilon} |f(x)| + \epsilon\|f\|, \qquad \textit{for all } f \in C_b(X). \tag{21}$$

Then there exists a unique finite (positive) Borel measure μ so that

$$\ell(f) = \int_X f(x)\, d\mu(x), \qquad \textit{for all } f \in C_b(X).$$

The extra hypothesis (21) (which is vacuous when X is compact) is a "tightness" assumption that will be relevant in Chapter 6. Note that as before $|\ell(f)| \leq \|f\|$ since $\ell(1) = 1$, even without the assumption (21).

The proof of this theorem proceeds as that of Theorem 7.1, save for one key aspect. First we define

$$\rho(\mathcal{O}) = \sup\{\ell(f),\ \text{where } f \in C_b(X),\ \operatorname{supp}(f) \subset \mathcal{O}, \text{ and } 0 \leq f \leq 1\}.$$

The change that is required is in the proof of the countable sub-additivity of ρ, in that the support of f's (in the definition of $\rho(\mathcal{O})$) are now not necessarily compact. In fact, suppose $\mathcal{O} = \bigcup_{k=1}^{\infty} \mathcal{O}_k$ is a countable union of open sets. Let C be the support of f, and given a fixed $\epsilon > 0$, set $K = C \cap K_\epsilon$, with K_ϵ the compact set arising in (21). Then K is compact and $\bigcup_{k=1}^{\infty} \mathcal{O}_k$ covers K. Proceeding as before, we obtain a partition of unity $\{\eta_k\}_{k=1}^N$, with η_k supported in $\mathcal{O}_k$ and $\sum_{k=1}^N \eta_k(x) = 1$, for $x \in K$. Now $f - \sum_{k=1}^N f\eta_k$ vanishes on K_ϵ. Thus by (21)

$$\left|\ell(f) - \sum_{k=1}^N \ell(f\eta_k)\right| \leq \epsilon,$$

and hence

$$\ell(f) \leq \sum_{k=1}^{\infty} \rho(\mathcal{O}_k) + \epsilon.$$

Since this holds for each ϵ, we obtain the required sub-additivity of ρ and thus of μ_*. The proof of the theorem can then be concluded as before.

Theorem 7.4 did not require that the metric space X be either complete or separable. However if we make these two further assumptions on X, then the condition (21) is actually necessary.

Indeed, suppose $\ell(f) = \int_X f\, d\mu$, where μ is a positive finite Borel measure on X, which we may assume is normalized, $\mu(X) = 1$. Under the assumption that X is complete and separable, then for each fixed $\epsilon > 0$ there is a compact set K_ϵ so that $\mu(K_\epsilon^c) < \epsilon$. Indeed, let $\{c_k\}$ be a dense sequence in X. Since for each m the collection of balls $\{B_{1/m}(c_k)\}_{k=1}^{\infty}$ covers X, there is a finite N_m so that if $\mathcal{O}_m = \bigcup_{k=1}^{N_m} B_{1/m}(c_k)$, then $\mu(\mathcal{O}_m) \geq 1 - \epsilon/2^m$.

Take $K_\epsilon = \bigcap_{m=1}^{\infty} \overline{\mathcal{O}_m}$. Then $\mu(K_\epsilon) \geq 1 - \epsilon$; also, K_ϵ is closed and totally bounded, in the sense that for every $\delta > 0$, the set K_ϵ can be covered by finitely many balls of radius δ. Since X is complete, K_ϵ must be compact. Now (21) follows immediately.

8 Exercises

1. Consider $L^p = L^p(\mathbb{R}^d)$ with Lebesgue measure. Let $f_0(x) = |x|^{-\alpha}$ if $|x| < 1$, $f_0(x) = 0$ for $|x| \geq 1$; also let $f_\infty(x) = |x|^{-\alpha}$ if $|x| \geq 1$, $f_\infty(x) = 0$ when $|x| < 1$. Show that:

(a) $f_0 \in L^p$ if and only if $p\alpha < d$.

(b) $f_\infty \in L^p$ if and only if $d < p\alpha$.

(c) What happens if in the definitions of f_0 and f_∞ we replace $|x|^{-\alpha}$ by $|x|^{-\alpha}/(\log(2/|x|))$ for $|x| < 1$, and $|x|^{-\alpha}$ by $|x|^{-\alpha}/(\log(2|x|))$ for $|x| \geq 1$?

2. Consider the spaces $L^p(\mathbb{R}^d)$, when $0 < p < \infty$.

(a) Show that if $\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$ for all f and g, then necessarily $p \geq 1$.

(b) Consider $L^p(\mathbb{R})$ where $0 < p < 1$. Show that there are no bounded linear functionals on this space. In other words, if ℓ is a linear function $L^p(\mathbb{R}) \to \mathbb{C}$ that satisfies

$$|\ell(f)| \leq M\|f\|_{L^p(\mathbb{R})} \qquad \text{for all } f \in L^p(\mathbb{R}) \text{ and some } M > 0,$$

then $\ell = 0$.

[Hint: For (a), prove that if $0 < p < 1$ and $x, y > 0$, then $x^p + y^p > (x+y)^p$. For (b), let F be defined by $F(x) = \ell(\chi_x)$, where χ_x is the characteristic function of $[0, x]$, and consider $F(x) - F(y)$.]

3. If $f \in L^p$ and $g \in L^q$, both not identically equal to zero, show that equality holds in Hölder's inequality (Theorem 1.1) if and only if there exist two non-zero constants $a, b \geq 0$ such that $a|f(x)|^p = b|g(x)|^q$ for a.e. x.

4. Suppose X is a measure space and $0 < p < 1$.

(a) Prove that $\|fg\|_{L^1} \geq \|f\|_{L^p}\|g\|_{L^q}$. Note that q, the conjugate exponent of p, is negative.

(b) Suppose f_1 and f_2 are non-negative. Then $\|f_1 + f_2\|_{L^p} \geq \|f_1\|_{L^p} + \|f_2\|_{L^p}$.

(c) The function $d(f, g) = \|f - g\|_{L^p}^p$ for $f, g \in L^p$ defines a metric on $L^p(X)$.

5. Let X be a measure space. Using the argument to prove the completeness of $L^p(X)$, show that if the sequence $\{f_n\}$ converges to f in the L^p norm, then a subsequence of $\{f_n\}$ converges to f almost everywhere.

6. Let $(X, \mathcal{F}, \mu)$ be a measure space. Show that:

(a) The simple functions are dense in $L^\infty(X)$ if $\mu(X) < \infty$, and;

(b) The simple functions are dense in $L^p(X)$ for $1 \leq p < \infty$.

[Hint: For (a), use $E_{\ell,j} = \{x \in X : \frac{M\ell}{j} \leq f(x) < \frac{M(\ell+1)}{j}\}$ where $-j \leq \ell \leq j$, and $M = \|f\|_{L^\infty}$. Then consider the functions f_j that equal $M\ell/j$ on $E_{\ell,j}$. For (b) use a construction similar to that in (a).]

7. Consider the L^p spaces, $1 \leq p < \infty$, on $\mathbb{R}^d$ with Lebesgue measure. Prove that:

(a) The family of continuous functions with compact support is dense in L^p, and in fact:

(b) The family of indefinitely differentiable functions with compact support is dense in L^p.

The cases of L^1 and L^2 are in Theorem 2.4, Chapter 2 of Book III, and Lemma 3.1, Chapter 5 of Book III.

8. Suppose $1 \leq p < \infty$, and that $\mathbb{R}^d$ is equipped with Lebesgue measure. Show that if $f \in L^p(\mathbb{R}^d)$, then

$$\|f(x+h) - f(x)\|_{L^p} \to 0 \qquad \text{as } |h| \to 0.$$

Prove that this fails when $p = \infty$.

[Hint: By the previous exercise, the continuous functions with compact support are dense in $L^p(\mathbb{R}^d)$ for $1 \leq p < \infty$. See also Theorem 2.4 and Proposition 2.5 in Chapter 2 of Book III.]

9. Suppose X is a measure space and $1 \leq p_0 < p_1 \leq \infty$.

(a) Consider $L^{p_0} \cap L^{p_1}$ equipped with

$$\|f\|_{L^{p_0} \cap L^{p_1}} = \|f\|_{L^{p_0}} + \|f\|_{L^{p_1}}.$$

Show that $\|\cdot\|_{L^{p_0} \cap L^{p_1}}$ is a norm, and that $L^{p_0} \cap L^{p_1}$ (with this norm) is a Banach space.

(b) Suppose $L^{p_0} + L^{p_1}$ is defined as the vector space of measurable functions f on X that can be written as a sum $f = f_0 + f_1$ with $f_0 \in L^{p_0}$ and $f_1 \in L^{p_1}$. Consider

$$\|f\|_{L^{p_0} + L^{p_1}} = \inf \left\{ \|f_0\|_{L^{p_0}} + \|f_1\|_{L^{p_1}} \right\},$$

where the infimum is taken over all decompositions $f = f_0 + f_1$ with $f_0 \in L^{p_0}$ and $f_1 \in L^{p_1}$. Show that $\|\cdot\|_{L^{p_0}+L^{p_1}}$ is a norm, and that $L^{p_0} + L^{p_1}$ (with this norm) is a Banach space.

(c) Show that $L^p \subset L^{p_0} + L^{p_1}$ if $p_0 \leq p \leq p_1$.

10. A measure space (X, μ) is **separable** if there is a countable family of measurable subsets $\{E_k\}_{k=1}^{\infty}$ so that if E is any measurable set of finite measure, then

$$\mu(E \triangle E_{n_k}) \to 0 \qquad \text{as } k \to 0$$

for an appropriate subsequence $\{n_k\}$ which depends on E. Here $A \triangle B$ denotes the symmetric difference of the sets A and B, that is,

$$A \triangle B = (A - B) \cup (B - A).$$

(a) Verify that $\mathbb{R}^d$ with the usual Lebesgue measure is separable.

(b) The space $L^p(X)$ is **separable** if there exists a countable collection of elements $\{f_n\}_{n=1}^{\infty}$ in L^p that is dense. Prove that if the measure space X is separable, then L^p is separable when $1 \leq p < \infty$.

11. In light of the previous exercise, prove the following:

(a) Show that the space $L^\infty(\mathbb{R})$ is not separable by constructing for each $a \in \mathbb{R}$ an $f_a \in L^\infty$, with $\|f_a - f_b\| \geq 1$, if $a = b$.

(b) Do the same for the dual space of $L^\infty(\mathbb{R})$.

12. Suppose the measure space (X, μ) is separable as defined in Exercise 10. Let $1 \leq p < \infty$ and $1/p + 1/q = 1$. A sequence $\{f_n\}$ with $f_n \in L^p$ is said to converge to $f \in L^p$ **weakly** if

$$\int f_n g \, d\mu \to \int fg \, d\mu \qquad \text{for every } g \in L^q. \tag{22}$$

(a) Verify that if $\|f - f_n\|_{L^p} \to 0$, then f_n converges to f weakly.

(b) Suppose $\sup_n \|f_n\|_{L^p} < \infty$. Then, to verify weak convergence it suffices to check (22) for a dense subset of functions g in L^q.

(c) Suppose $1 < p < \infty$. Show that if $\sup_n \|f_n\|_{L^p} < \infty$, then there exists $f \in L^p$, and a subsequence $\{n_k\}$ so that f_{n_k} converges weakly to f.

Part (c) is known as the "weak compactness" of L^p for $1 < p < \infty$, which fails when $p = 1$ as is seen in the exercise below.

[Hint: For (b) use Exercise 10 (b).]

13. Below are some examples illustrating weak convergence.

(a) $f_n(x) = \sin(2\pi n x)$ in $L^p([0,1])$. Show that $f_n \to 0$ weakly.

(b) $f_n(x) = n^{1/p}\chi(nx)$ in $L^p(\mathbb{R})$. Then $f_n \to 0$ weakly if $p > 1$, but not when $p = 1$. Here χ denotes the characteristic function of $[0,1]$.

(c) $f_n(x) = 1 + \sin(2\pi n x)$ in $L^1([0,1])$. Then $f_n \to 1$ weakly also in $L^1([0,1])$, $\|f_n\|_{L^1} = 1$, but $\|f_n - 1\|_{L^1}$ does not converge to zero. Compare with Problem 6 part (d).

14. Suppose X is a measure space, $1 < p < \infty$, and suppose $\{f_n\}$ is a sequence of functions with $\|f_n\|_{L^p} \leq M < \infty$.

(a) Prove that if $f_n \to f$ a.e. then $f_n \to f$ weakly.

(b) Show that the above result may fail if $p = 1$.

(c) Show that if $f_n \to f_1$ a.e. and $f_n \to f_2$ weakly, then $f_1 = f_2$ a.e.

15. Minkowski's inequality for integrals. Suppose (X_1, μ_1) and (X_2, μ_2) are two measure spaces, and $1 \leq p \leq \infty$. Show that if $f(x_1, x_2)$ is measurable on $X_1 \times X_2$ and non-negative, then

$$\left\| \int f(x_1, x_2) \, d\mu_2 \right\|_{L^p(X_1)} \leq \int \|f(x_1, x_2)\|_{L^p(X_1)} \, d\mu_2.$$

Extend this statement to the case when f is complex-valued and the right-hand side of the inequality is finite.

[Hint: For $1 < p < \infty$, use a combination of Hölder's inequality, and its converse in Lemma 4.2.]

16. Prove that if $f_j \in L^{p_j}(X)$, where X is a measure space, $j = 1, \ldots, N$, and $\sum_{j=1}^N 1/p_j = 1$ with $p_j \geq 1$, then

$$\|\prod_{j=1}^N f_j\|_{L^1} \leq \prod_{j=1}^N \|f_j\|_{L^{p_j}}.$$

This is the multiple Hölder inequality.

17. The **convolution** of f and g on $\mathbb{R}^d$ equipped with the Lebesgue measure is defined by

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

(a) If $f \in L^p$, $1 \leq p \leq \infty$, and $g \in L^1$, then show that for almost every x the integrand $f(x-y)g(y)$ is integrable in y, hence $f * g$ is well defined. Moreover, $f * g \in L^p$ with

$$\|f * g\|_{L^p} \leq \|f\|_{L^p}\|g\|_{L^1}.$$

(b) A version of (a) applies when g is replaced by a finite Borel measure μ: if $f \in L^p$, with $1 \leq p \leq \infty$, define

$$(f * \mu)(x) = \int_{\mathbb{R}^d} f(x-y)\,d\mu(y),$$

and show that $\|f * \mu\|_{L^p} \leq \|f\|_{L^p}|\mu|(\mathbb{R}^d)$.

(c) Prove that if $f \in L^p$ and $g \in L^q$, where p and q are conjugate exponents, then $f * g \in L^\infty$ with $\|f * g\|_{L^\infty} \leq \|f\|_{L^p}\|g\|_{L^q}$. Moreover, the convolution $f * g$ is uniformly continuous on $\mathbb{R}$, and if $1 < p < \infty$, then $\lim_{|x|\to\infty}(f * g)(x) = 0$.

[Hint: For (a) and (b) use the Minkowski inequality for integrals in Exercise 15. For part (c), use Exercise 8.]

18. We consider the L^p spaces with **mixed norm**, in a special case that is useful is several contexts.

We take as our underlying space the product space $\{(x,t)\} = \mathbb{R}^d \times \mathbb{R}$, with the product measure $dx\,dt$, where dx and dt are Lebesgue measures on $\mathbb{R}^d$ and $\mathbb{R}$ respectively. We define $L_t^r(L_x^p) = L^{p,r}$, with $1 \leq p \leq \infty$, $1 \leq r \leq \infty$, to be the space of equivalence classes of jointly measurable functions $f(x,t)$ for which the norm

$$\|f\|_{L^{p,r}} = \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^d} |f(x,t)|^p \, dx\right)^{\frac{r}{p}} dt\right)^{\frac{1}{r}}$$

is finite (when $p < \infty$ and $r < \infty$), and an obvious variant when $p = \infty$ or $r = \infty$.

(a) Verify that $L^{p,r}$ with this norm is complete, and hence is a Banach space.

(b) Prove the general form of Hölder's inequality in this context

$$\iint_{\mathbb{R}^d \times \mathbb{R}} |f(x,t)g(x,t)| \, dx \, dt \leq \|f\|_{L^{p,r}} \|g\|_{L^{p',r'}},$$

with $1/p + 1/p' = 1$ and $1/r + 1/r' = 1$.

(c) Show that if f is integrable over all sets of finite measure, then

$$\|f\|_{L^{p,r}} = \sup \left| \int_{\mathbb{R}^d \times \mathbb{R}} f(x,t)g(x,t) \, dx dt \right|,$$

with the sup taken over all g that are simple and $\|g\|_{L^{p',r'}} \leq 1$.

(d) Conclude that the dual space of $L^{p,r}$ is $L^{p',r'}$, if $1 \leq p < \infty$, and $1 \leq r < \infty$.

19. Young's inequality. Suppose $1 \leq p, q, r \leq \infty$. Prove the following on $\mathbb{R}^d$:

$$\|f * g\|_{L^q} \leq \|f\|_{L^p} \|g\|_{L^r} \qquad \text{whenever } 1/q = 1/p + 1/r - 1.$$

Here, $f * g$ denotes the convolution of f and g as defined in Exercise 17.
[Hint: Assume $f, g \geq 0$, and use the decomposition

$$f(y)g(x-y) = f(y)^a g(x-y)^b [f(y)^{1-a} g(x-y)^{1-b}]$$

for appropriate a and b, together with Exercise 16 to find that

$$\left| \int f(y)g(x-y) \, dy \right| \leq \|f\|_{L^p}^{1-q/p} \|g\|_{L^q}^{1-q/r} \left(\int |f(y)|^p |g(x-y)|^r \, dy \right)^{\frac{1}{q}}.]$$

20. Suppose X is a measure space, $0 < p_0 < p < p_1 \leq \infty$, and $f \in L^{p_0}(X) \cap L^{p_1}(X)$. Then $f \in L^p(X)$ and

$$\|f\|_{L^p} \leq \|f\|_{L^{p_0}}^{1-t} \|f\|_{L^{p_1}}^{t}, \qquad \text{if } t \text{ is chosen so that } \frac{1}{p} = \frac{1-t}{p_0} + \frac{t}{p_1}.$$

21. Recall the definition of a convex function. (See Problem 4, Chapter 3, in Book III.) Suppose φ is a non-negative convex function on $\mathbb{R}$ and f is real-valued and integrable on a measure space X, with $\mu(X) = 1$. Then we have **Jensen's inequality**:

$$\varphi\left(\int_X f\,d\mu\right) \leq \int_X \varphi(f)\,d\mu.$$

Note that if $\varphi(t) = |t|^p$, $1 \leq p$, then φ is convex and the above can be obtained from Hölder's inequality. Another interesting case is $\varphi(t) = e^{at}$.

[Hint: Since φ is convex, one has, $\varphi(\sum_{j=1}^N a_j x_j) \leq \sum_{j=1}^N a_j \varphi(x_j)$, whenever a_j, x_j are real, $a_j \geq 0$, and $\sum_{j=1}^N a_j = 1$.]

22. Another inequality of Young. Suppose φ and ψ are both continuous, *strictly* increasing functions on $[0, \infty)$ that are inverses of each other, that is, $(\varphi \circ \psi)(x) = x$ for all $x \geq 0$. Let

$$\Phi(x) = \int_0^x \varphi(u)\,du \quad \text{and} \quad \Psi(x) = \int_0^x \psi(u)\,du.$$

(a) Prove: $ab \leq \Phi(a) + \Psi(b)$ for all $a, b \geq 0$.

In particular, if $\varphi(x) = x^{p-1}$ and $\psi(y) = y^{q-1}$ with $1 < p < \infty$ and $1/p + 1/q = 1$, then we get $\Phi(x) = x^p/p$, $\Psi(y) = y^q/q$, and

$$A^\theta B^{1-\theta} \leq \theta A + (1-\theta)B \qquad \text{for all } A, B \geq 0 \text{ and } 0 \leq \theta \leq 1.$$

(b) Prove that we have equality in Young's inequality only if $b = \varphi(a)$ (that is, $a = \psi(b)$).

[Hint: Consider the area ab of the rectangle whose vertices are $(0,0)$, $(a,0)$, $(0,b)$ and (a,b), and compare it to areas "under" the curves $y = \Phi(x)$ and $x = \Psi(y)$.]

23. Let (X, μ) be a measure space and suppose $\Phi(t)$ is a continuous, convex, and increasing function on $[0, \infty)$, with $\Phi(0) = 0$. Define

$$L^\Phi = \{f \text{ measurable} : \int_X \Phi(|f(x)|/M)\,d\mu < \infty \text{ for some } M > 0\},$$

and

$$\|f\|_\Phi = \inf_{M>0} \int_X \Phi(|f(x)|/M)\,d\mu \leq 1.$$

Prove that:

(a) L^Φ is a vector space.

(b) $\|\cdot\|_{L^\Phi}$ is a norm.

(c) L^Φ is complete in this norm.

The Banach spaces L^Φ are called **Orlicz spaces**. Note that in the special case $\Phi(t) = t^p$, $1 \le p < \infty$, then $L^\Phi = L^p$.

[Hint: Observe that if $f \in L^\Phi$, then $\lim_{N\to\infty} \int_X \Phi(|f|/N)\, d\mu = 0$. Also, use the fact that there exists $A > 0$ so that $\Phi(t) \ge At$ for all $t \ge 0$.]

24. Let $1 \le p_0 < p_1 < \infty$.

(a) Consider the Banach space $L^{p_0} \cap L^{p_1}$ with norm $\|f\|_{L^{p_0}\cap L^{p_1}} = \|f\|_{L^{p_0}} + \|f\|_{L^{p_1}}$. (See Exercise 9.) Let

$$\Phi(t) = \begin{cases} t^{p_0} & \text{if } 0 \le t \le 1, \\ t^{p_1} & \text{if } 1 \le t < \infty. \end{cases}$$

Show that L^Φ with its norm is equivalent to the space $L^{p_0} \cap L^{p_1}$. In other words, there exist $A, B > 0$, so that

$$A\|f\|_{L^{p_0}\cap L^{p_1}} \le \|f\|_{L^\Phi} \le B\|f\|_{L^{p_0}\cap L^{p_1}}.$$

(b) Similarly, consider the Banach space $L^{p_0} + L^{p_1}$ with its norm as defined in Exercise 9. Let

$$\Psi(t) = \int_0^t \psi(u)\, du \qquad \text{where} \qquad \psi(u) = \begin{cases} u^{p_1-1} & \text{if } 0 \le u \le 1, \\ u^{p_0-1} & \text{if } 1 \le u < \infty. \end{cases}$$

Show that L^Ψ with its norm is equivalent to the space $L^{p_0} + L^{p_1}$.

25. Show that a Banach space $\mathcal{B}$ is a Hilbert space if and only if the parallelogram law holds

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|^2 + \|g\|^2).$$

As a consequence, prove that if $L^p(\mathbb{R}^d)$ with the Lebesgue measure is a Hilbert space, then necessarily $p = 2$.

[Hint: For the first part, in the real case, let $(f,g) = \frac{1}{4}(\|f+g\|^2 + \|f-g\|^2)$.]

26. Suppose $1 < p_0, p_1 < \infty$ and $1/p_0 + 1/q_0 = 1$ and $1/p_1 + 1/q_1 = 1$. Show that the Banach spaces $L^{p_0} \cap L^{p_1}$ and $L^{q_0} + L^{q_1}$ are duals of each other up to an equivalence of norms. (See Exercise 9 for the relevant definitions of these spaces. Also, Problem 5* gives a generalization of this result.)

27. The purpose of this exercise is to prove that the unit ball in L^p is strictly convex when $1 < p < \infty$, in the following sense. Here L^p is the space of real-valued functions whose p^{th} power are integrable. Suppose $\|f_0\|_{L^p} = \|f_1\|_{L^p} = 1$, and let

$$f_t = (1-t)f_0 + tf_1$$

be the straight-line segment joining the points f_0 and f_1. Then $\|f_t\|_{L^p} < 1$ for all t with $0 < t < 1$, unless $f_0 = f_1$.

(a) Let $f \in L^p$ and $g \in L^q$, $1/p + 1/q = 1$, with $\|f\|_{L^p} = 1$ and $\|g\|_{L^q} = 1$. Then

$$\int fg \, d\mu = 1$$

only when $f(x) = \operatorname{sign} g(x) |g(x)|^{q-1}$.

(b) Suppose $\|f_{t'}\|_{L^p} = 1$ for some $0 < t' < 1$. Find $g \in L^q$, $\|g\|_{L^q} = 1$, so that

$$\int f_{t'} g \, d\mu = 1$$

and let $F(t) = \int f_t g \, d\mu$. Observe as a result that $F(t) = 1$ for all $0 \leq t \leq 1$. Conclude that $f_t = f_0$ for all $0 \leq t \leq 1$.

(c) Show that the strict convexity fails when $p = 1$ or $p = \infty$. What can be said about these cases?

A stronger assertion is given in Problem 6*.
[Hint: To prove (a) show that the case of equality in $A^\theta B^{1-\theta} \leq \theta A + (1 - \theta)B$, for $A, B > 0$ and $0 < \theta < 1$ holds only when $A = B$.]

28. Verify the completeness of $\Lambda^\alpha(\mathbb{R}^d)$ and $L^p_k(\mathbb{R}^d)$.

29. Consider further the spaces $\Lambda^\alpha(\mathbb{R}^d)$.

(a) Show that when $\alpha > 1$ the only functions in $\Lambda^\alpha(\mathbb{R}^d)$ are the constants.

(b) Motivated by (a), one defines $C^{k,\alpha}(\mathbb{R}^d)$ to be the class of functions f on $\mathbb{R}^d$ whose partial derivatives of order less than or equal to k belong to $\Lambda^\alpha(\mathbb{R}^d)$. Here k is an integer and $0 < \alpha \leq 1$. Show that this space, endowed with the norm

$$\|f\|_{C^{k,\alpha}} = \sum_{|\beta| \leq k} \left\| \partial_x^\beta f \right\|_{\Lambda^\alpha(\mathbb{R}^d)},$$

is a Banach space.

30. Suppose $\mathcal{B}$ is a Banach space and $\mathcal{S}$ is a closed linear subspace of $\mathcal{B}$. The subspace $\mathcal{S}$ defines an equivalence relation $f \sim g$ to mean $f - g \in \mathcal{S}$. If $\mathcal{B}/\mathcal{S}$ denotes the collection of these equivalence classes, then show that $\mathcal{B}/\mathcal{S}$ is a Banach space with norm $\|f\|_{\mathcal{B}/\mathcal{S}} = \inf(\|f'\|_{\mathcal{B}},\ \ f' \sim f)$.

31. If Ω is an open subset of $\mathbb{R}^d$ then one definition of $L^p_k(\Omega)$ can be taken to be the quotient Banach space $\mathcal{B}/\mathcal{S}$, as defined in the previous exercise, with $\mathcal{B} = L^p_k(\mathbb{R}^d)$ and $\mathcal{S}$ the subspace of those functions which vanish a.e. on Ω. Another possible space, that we will denote by $L^p_k(\Omega^0)$, consists of the closure in $L^p_k(\mathbb{R}^d)$ of all f that have compact support in Ω. Observe that the natural mapping of $L^p_k(\Omega^0)$ to

$L_k^p(\Omega)$ has norm equal to 1. However, this mapping is in general not surjective. Prove this in the case when Ω is the unit ball and $k \geq 1$.

32. A Banach space is said to be separable if it contains a countable dense subset. In Exercise 11 we saw an example of a Banach space $\mathcal{B}$ that is separable, but where $\mathcal{B}^*$ is not separable. Prove, however, that in general when $\mathcal{B}^*$ is separable, then $\mathcal{B}$ is separable. Note that this gives another proof that in general L^1 is not the dual of L^∞.

33. Let V be a vector space over the complex numbers $\mathbb{C}$, and suppose there exists a real-valued function p on V satisfying:

$$\begin{cases} p(\alpha v) = |\alpha| p(v), & \text{if } \alpha \in \mathbb{C}, \text{ and } v \in V, \\ p(v_1 + v_2) \leq p(v_1) + p(v_2), & \text{if } v_1 \text{ and } v_2 \in V. \end{cases}$$

Prove that if V_0 is a subspace of V and ℓ_0 a linear functional on V_0 which satisfies $|\ell_0(f)| \leq p(f)$ for all $f \in V_0$, then ℓ_0 can be extended to a linear functional ℓ on V that satisfies $|\ell(f)| \leq p(f)$ for all $f \in V$.

[Hint: If $u = \mathrm{Re}(\ell_0)$, then $\ell_0(v) = u(v) - iu(iv)$. Apply Theorem 5.2 to u.]

34. Suppose $\mathcal{B}$ is a Banach space and $\mathcal{S}$ a closed proper subspace, and assume $f_0 \notin \mathcal{S}$. Show that there is a continuous linear functional ℓ on $\mathcal{B}$, so that $\ell(f) = 0$ for $f \in \mathcal{S}$, and $\ell(f_0) = 1$. The linear functional ℓ can be chosen so that $\|\ell\| = 1/d$ where d is the distance from f_0 to $\mathcal{S}$.

35. A linear functional ℓ on a Banach space $\mathcal{B}$ is continuous if and only if $\{f \in \mathcal{B} : \ell(f) = 0\}$ is closed.

[Hint: This is a consequence of Exercise 34.]

36. The results in Section 5.4 can be extended to d-dimensions.

(a) Show that there exists an extended-valued non-negative function $\hat{m}$ defined on all subsets of $\mathbb{R}^d$ so that (i) $\hat{m}$ is finitely additive; (ii) $\hat{m}(E) = m(E)$ whenever E is Lebesgue measurable, where m is Lebesgue measure; and $\hat{m}(E+h) = \hat{m}(E)$ for all sets E and every $h \in \mathbb{R}^d$. Prove this is as a consequence of (b) below.

(b) Show that there is an "integral" I, defined on all bounded functions on $\mathbb{R}^d/\mathbb{Z}^d$, so that $I(f) \geq 0$ whenever $f \geq 0$; the map $f \mapsto I(f)$ is linear; $I(f) = \int_{\mathbb{R}^d/\mathbb{Z}^d} f\, dx$ whenever f is measurable; and $I(f_h) = I(f)$ where $f_h(x) = f(x - h)$, and $h \in \mathbb{R}^d$.

9 Problems

1. The spaces L^∞ and L^1 play universal roles with respect to all Banach spaces in the following sense.

(a) If $\mathcal{B}$ is any separable Banach space, show that it can be realized without change of norm as a linear subspace of $L^\infty(\mathbb{Z})$. Precisely, prove that there is a linear operator i of $\mathcal{B}$ into $L^\infty(\mathbb{Z})$ so that $\|i(f)\|_{L^\infty(\mathbb{Z})} = \|f\|_{\mathcal{B}}$ for all $f \in \mathcal{B}$.

(b) Each such $\mathcal{B}$ can also be realized as a quotient space of $L^1(\mathbb{Z})$. That is, there is a linear surjection P of $L^1(\mathbb{Z})$ onto $\mathcal{B}$, so that if $\mathcal{S} = \{x \in L^1(\mathbb{Z}) :\ P(x) = 0\}$, then $\|P(x)\|_{\mathcal{B}} = \inf_{y\in\mathcal{S}} \|x+y\|_{L^1(\mathbb{Z})}$, for each $x \in L^1(\mathbb{Z})$. This gives an identification of $\mathcal{B}$ (and its norm) with the quotient space $L^1(\mathbb{Z})/\mathcal{S}$ (and its norm), as defined in Exercise 30.

Note that similar conclusions hold for $L^\infty(X)$ and $L^1(X)$ if X is a measure space that contains a countable disjoint collection of measurable sets of positive and finite measure.

[Hint: For (a), let $\{f_n\}$ be a dense set of non-zero vectors in $\mathcal{B}$, and let $\ell_n \in \mathcal{B}^*$ be such that $\|\ell_n\|_{\mathcal{B}^*} = 1$ and $\ell_n(f_n) = \|f_n\|$. If $f \in \mathcal{B}$, set $i(f) = \{\ell_n(f)\}_{-\infty}^{\infty}$. For (b), if $x = \{x_n\} \in L^1(\mathbb{Z})$, with $\sum_{-\infty}^{\infty} |x_n| = \|x\|_{L^1(\mathbb{Z})} < \infty$, define P by $P(x) = \sum_{-\infty}^{\infty} x_n f_n/\|f_n\|$.]

2. There is a "generalized limit" L defined on the vector space V of all real sequences $\{s_n\}_{n=1}^{\infty}$ that are bounded, so that:

(i) L is a linear functional on V.

(ii) $L(\{s_n\}) \geq 0$ if $s_n \geq 0$, for all n.

(iii) $L(\{s_n\}) = \lim_{n\to\infty} s_n$ if the sequence $\{s_n\}$ has a limit.

(iii) $L(\{s_n\}) = L(\{s_{n+k}\})$ for every $k \geq 1$.

(iii) $L(\{s_n\}) = L(\{s_{n'}\})$ if $s_n - s_n' = 0$ for only finitely many n.

[Hint: Let $p(\{s_n\}) = \limsup_{n\to\infty} \left(\frac{s_1+\cdots+s_n}{n}\right)$, and extend the linear functional L defined by $L(\{s_n\}) = \lim_{n\to\infty} s_n$, defined on the subspace consisting of sequences that have limits.]

3. Show that the closed unit ball in a Banach space $\mathcal{B}$ is compact (that is, if $f_n \in \mathcal{B}$, $\|f_n\| \leq 1$, then there is a subsequence that converges in the norm) if and only if $\mathcal{B}$ is finite dimensional.

[Hint: If $\mathcal{S}$ is a closed subspace of $\mathcal{B}$, then there exists $x \in \mathcal{B}$ with $\|x\| = 1$ and the distance between x and $\mathcal{S}$ is greater than $1/2$.]

4. Suppose X is a σ-compact measurable metric space, and $C_b(X)$ is separable, where $C_b(X)$ denotes the Banach space of bounded continuous functions on X with the sup-norm.

(a) If $\{\mu_n\}_{n=1}^{\infty}$ is a bounded sequence in $M(X)$, then there exists a $\mu \in M(X)$ and a subsequence $\{\mu_{n_j}\}_{j=1}^{\infty}$, so that μ_{n_j} converges to μ in the following (weak*) sense:

$$\int_X g(x)\,d\mu_{n_j}(x) \to \int_X g(x)\,d\mu(x), \qquad \text{for all } g \in C_b(X).$$

(b) Start with a $\mu_0 \in M(X)$ that is positive, and for each $f \in L^1(\mu_0)$ consider the mapping $f \mapsto f d\mu_0$. This mapping is an isometry of $L^1(\mu_0)$ to the subspace of $M(X)$ consisting of signed measures which are absolutely continuous with respect to μ_0.

(c) Hence if $\{f_n\}$ is a bounded sequence of functions in $L^1(\mu_0)$, then there exist a $\mu \in M(X)$ and a subsequence $\{f_{n_j}\}$ such that the measures $f_{n_j} d\mu_0$ converge to μ in the above sense.

5.* Let X be a measure space. Suppose φ and ψ are both continuous, strictly increasing functions on $[0, \infty)$ which are inverses of each other, that is, $(\varphi \circ \psi)(x) = x$ for all $x \geq 0$. Let

$$\Phi(x) = \int_0^x \varphi(u)\, du \qquad \text{and} \qquad \Psi(x) = \int_0^x \psi(u)\, du.$$

Consider the Orlicz spaces $L^\Phi(X)$ and $L^\Psi(X)$ introduced in Exercise 23.

(a) In connection with Exercise 22 the following Hölder-like inequality holds:

$$\int |fg| \leq C\|f\|_{L^\Phi} \|g\|_{L^\Psi} \qquad \text{for some } C > 0, \text{ and all } f \in L^\Phi \text{ and } g \in L^\Psi.$$

(b) Suppose there exists $c > 0$ so that $\Phi(2t) \leq c\Phi(t)$ for all $t \geq 0$. Then the dual of L^Φ is equivalent to L^Ψ.

6.* There are generalizations of the parallelogram law for L^2 (see Exercise 25) that hold for L^p. These are the Clarkson inequalities:

(a) For $2 \leq p \leq \infty$ the statement is that

$$\left\| \frac{f+g}{2} \right\|_{L^p}^p + \left\| \frac{f-g}{2} \right\|_{L^p}^p \leq \frac{1}{2} \left(\|f\|_{L^p}^p + \|g\|_{L^p}^p \right).$$

(b) For $1 < p \leq 2$ the statement is that

$$\left\| \frac{f+g}{2} \right\|_{L^p}^q + \left\| \frac{f-g}{2} \right\|_{L^p}^q \leq \frac{1}{2} \left(\|f\|_{L^p}^p + \|g\|_{L^p}^p \right)^{q/p},$$

where $1/p + 1/q = 1$.

(c) As a result, L^p is **uniformly convex** when $1 < p < \infty$. This means that there is a function $\delta = \delta(\epsilon) = \delta_p(\epsilon)$, with $0 < \delta < 1$, (and $\delta(\epsilon) \to 0$ as $\epsilon \to 0$), so that whenever $\|f\|_{L^p} = \|g\|_{L^p} = 1$, then $\|f - g\|_{L^p} \geq \epsilon$ implies that $\|\frac{f+g}{2}\| \leq 1 - \delta$.

This is stronger than the conclusion of strict convexity in Exercise 27.

(d) Using the result in (c), prove the following: suppose $1 < p < \infty$, and the sequence $\{f_n\}$, $f_n \in L^p$, converges weakly to f. If $\|f_n\|_{L^p} \to \|f\|_{L^p}$, then f_n converges to f strongly, that is, $\|f_n - f\|_{L^p} \to 0$ as $n \to \infty$.

7.* An important notion is that of the equivalence of Banach spaces. Suppose $\mathcal{B}_1$ and $\mathcal{B}_2$ are a pair of Banach spaces. We say that $\mathcal{B}_1$ and $\mathcal{B}_2$ are **equivalent** (also said to be "isomorphic") if there is a linear bijection T between $\mathcal{B}_1$ and $\mathcal{B}_2$ that is bounded and whose inverse is also bounded. Note that any pair of finite-dimensional Banach spaces are equivalent if and only if their dimensions are the same.

Suppose now we consider $L^p(X)$ for a general class of X (which contains for instance, $X = \mathbb{R}^d$ with Lebesgue measure). Then:

(a) L^p and L^q are equivalent if and only if $p = q$.

(b) However, for any p with $1 \leq p \leq \infty$, L^2 is equivalent with a closed infinite-dimensional subspace of L^p.

8.* There is no finitely-additive rotationally-invariant measure extending Lebesgue measure to all subsets of the sphere S^d when $d \geq 2$, in distinction to what happens on the torus $\mathbb{R}^d/\mathbb{Z}^d$ when $d \geq 2$. (See Exercise 36). This is due to a remarkable construction of Hausdorff that uses the fact that the corresponding rotation group of S^d is non-commutative. In fact, one can decompose S^2 into four disjoint sets A, B, C and Z so that (i) Z is denumerable, (ii) $A \sim B \sim C$, but $A \sim (B \cup C)$.

Here the notation $A_1 \sim A_2$ means that A_1 can be transformed into A_2 via a rotation.

9.* As a consequence of the previous problem one can show that it is not possible to extend Lebesgue measure on $\mathbb{R}^d$, $d \geq 3$, as a finitely-additive measure on all subsets of $\mathbb{R}^d$ so that it is both translation and rotation invariant (that is, invariant under Euclidean motions). This is graphically shown by the "Banach-Tarski paradox": There is a finite decomposition of the unit ball $B_1 = \bigcup_{j=1}^N E_j$, with the sets E_j disjoint, and there are corresponding sets $\tilde{E}_j$ that are each obtained from E_j by a Euclidean motion, with the $\tilde{E}_j$ also disjoint, so that $\bigcup_{j=1}^N \tilde{E}_j = B_2$ the ball of radius 2.