1.19 Approximation by C^∞ functions 36

Exercises 37

CHAPTER 2. L^p-Spaces 41

2.1 Definition of L^p-spaces 41
2.2 Jensen's inequality 44
2.3 Hölder's inequality 45
2.4 Minkowski's inequality 47
2.5 Hanner's inequality 49
2.6 Differentiability of norms 51
2.7 Completeness of L^p-spaces 52
2.8 Projection on convex sets 53
2.9 Continuous linear functionals and weak convergence 54
2.10 Linear functionals separate 56
2.11 Lower semicontinuity of norms 57
2.12 Uniform boundedness principle 58
2.13 Strongly convergent convex combinations 60
2.14 The dual of $L^p(\Omega)$ 61
2.15 Convolution 64
2.16 Approximation by C^∞-functions 64
2.17 Separability of $L^p(\mathbb{R}^n)$ 67
2.18 Bounded sequences have weak limits 68
2.19 Approximation by C_c^∞-functions 69
2.20 Convolutions of functions in dual $L^p(\mathbb{R}^n)$-spaces are continuous 70
2.21 Hilbert-spaces 71

Exercises 75

CHAPTER 3. Rearrangement Inequalities 79

3.1 Introduction 79
3.2 Definition of functions vanishing at infinity 80
3.3 Rearrangements of sets and functions 80
3.4 The simplest rearrangement inequality 82
3.5 Nonexpansivity of rearrangement 83

L^p-Spaces

This and the next two chapters contain basic facts about functions, the objects of principal interest in the rest of the book. The main topic is the definition and properties of p^{th}-power summable functions.

This topic does not utilize any metric properties of the domain, e.g., the Euclidean structure of $\mathbb{R}^n$, and therefore can be stated in greater generality than we shall actually need later. This generality is sometimes useful in other contexts, however. On a first reading it may be simplest to replace the measure $\mu(\mathrm{d}x)$ on the space Ω by Lebesgue measure $\mathrm{d}x$ on $\mathbb{R}^n$ and to regard Ω as a Lebesgue measurable subset of $\mathbb{R}^n$.

2.1 DEFINITION OF L^p-SPACES

Let Ω be a measure space with a (positive) measure μ and let $1 \leq p < \infty$. We define $L^p(\Omega, \mathrm{d}\mu)$ to be the following class of measurable functions:

$$L^p(\Omega, \mathrm{d}\mu) = \{f \ : \ f : \Omega \to \mathbb{C},\ f \text{ is } \mu\text{-measurable and } |f|^p \text{ is } \mu\text{-summable}\}. \tag{1}$$

Usually we omit μ in the notation and write instead $L^p(\Omega)$ if there is no ambiguity. Most of the time we have in mind that Ω is a Lebesgue measurable subset of $\mathbb{R}^n$ and μ is Lebesgue measure.

The reason we exclude $p < 1$ is that 3(c) below fails when $p < 1$.

On account of the inequality $|\alpha + \beta|^p \leq 2^{p-1}(|\alpha|^p + |\beta|^p)$ we see that for arbitrary complex numbers a and b, $af + bg$ is in $L^p(\Omega)$ if f and g are. Thus $L^p(\Omega)$ is a vector space.

For each $f \in L^p(\Omega)$ we define the **norm** to be

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \mu(\mathrm{d}x)\right)^{1/p}. \tag{2}$$

Sometimes we shall write this as $\|f\|_{L^p(\Omega)}$ if there is possibility of confusion. This norm has the following *three crucial properties* that make it truly a norm:

$$\begin{aligned}
&\text{(a)} \quad \|\lambda f\|_p = |\lambda| \|f\|_p \text{ for } \lambda \in \mathbb{C}. \\
&\text{(b)} \quad \|f\|_p = 0 \text{ if and only if } f(x) = 0 \text{ for } \mu\text{-almost every point } x. \\
&\text{(c)} \quad \|f+g\|_p \leq \|f\|_p + \|g\|_p.
\end{aligned} \tag{3}$$

(Technically, (2) only defines a semi-norm because of the 'almost every' caveat in 3(b), i.e., $\|f\|_p$ can be zero without $f \equiv 0$. Later on, when we define equivalence classes, (2) will be an honest norm on these classes.) Property (a) is obvious and (b) follows from the definition of the integral. Less trivial is property (c) which is called the **triangle inequality**. It will follow immediately from Theorem 2.4 (Minkowski's inequality). The triangle inequality is the same thing as **convexity of the norm**, i.e., if $0 \leq \lambda \leq 1$, then

$$\|\lambda f + (1-\lambda)g\|_p \leq \lambda \|f\|_p + (1-\lambda)\|g\|_p.$$

We can also define $L^\infty(\Omega, \mathrm{d}\mu)$ by

$$\begin{aligned}
L^\infty(\Omega, \mathrm{d}\mu) = \{f \ : \ f : \Omega \to \mathbb{C},\ f \text{ is } \mu\text{-measurable and there exists} \\
\text{a finite constant } K \text{ such that } |f(x)| \leq K \ \text{ for } \ \mu\text{-a.e. } x \in \Omega\}.
\end{aligned} \tag{4}$$

For $f \in L^\infty(\Omega)$ we define the norm

$$\|f\|_\infty = \inf\{K : \ |f(x)| \leq K \text{ for } \mu\text{-almost every } x \in \Omega\}. \tag{5}$$

Note that the norm depends on μ. This quantity is also called the **essential supremum** of $|f|$ and is denoted by ess $\sup_x |f(x)|$. (Do not confuse this with ess supp—which has one more p.) Unlike the usual supremum, ess sup ignores sets of μ-measure zero. E.g., if $\Omega = \mathbb{R}$ and $f(x) = 1$ if x is rational and $f(x) = 0$ otherwise, then (with respect to Lebesgue measure) ess $\sup_x |f(x)| = 0$, while $\sup_x |f(x)| = 1$.

One can easily verify that the L^∞ norm has the same properties (a), (b) and (c) as above. Note that property (b) would fail if ess sup is replaced by sup. Also note that $|f(x)| \leq \|f\|_\infty$ for almost every x.

We leave it as an exercise to the reader to prove that when $f \in L^\infty(\Omega) \cap L^q(\Omega)$ for some q then $f \in L^p(\Omega)$ for all $p > q$ and

$$\|f\|_\infty = \lim_{p\to\infty} \|f\|_p. \tag{6}$$

This equation is the reason for denoting the space defined in (4) by $L^\infty(\Omega)$.

An important concept, whose meaning will become clear later, is the **dual index** to p (for $1 \le p \le \infty$, of course). This is often denoted by p', but we shall often use q, and it is given by

$$\frac{1}{p} + \frac{1}{p'} = 1. \tag{7}$$

Thus, 1 and ∞ are dual, while the dual of 2 is 2.

Unfortunately, the norms we have defined do not serve to distinguish all different measurable functions, i.e., if $\|f - g\|_p = 0$ we can only conclude that $f(x) = g(x)$ μ-almost everywhere. To deal with this nuisance we can redefine $L^p(\Omega, \mathrm{d}\mu)$ so that its elements are not functions but *equivalence classes* of functions. That is to say, if we pick an $f \in L^p(\Omega)$ we can define $\widetilde{f}$ to be the set of all those functions that differ from f only on a set of μ-measure zero. If h is such a function we write $f \sim h$; moreover if $f \sim h$ and $h \sim g$, then $f \sim g$. Consequently, two such sets $\widetilde{f}$ and $\widetilde{k}$ are either identical or disjoint. We can now define

$$\|\widetilde{f}\|_p := \|f\|_p$$

for some $f \in \widetilde{f}$. The point is that this definition does not depend on the choice of $f \in \widetilde{f}$.

Thus we have two vector spaces. The first consists of functions while the second consists of equivalence classes of functions. (It is left to the reader to understand how to make the set of equivalence classes into a vector space.) For the first, $\|f - g\|_p = 0$ does not imply $f = g$, but for the second space it does. Some authors distinguish these spaces by different symbols, but all agree that it is the second space that should be called $L^p(\Omega)$. Nevertheless most authors will eventually slip into the tempting trap of saying 'let f be a function in $L^p(\Omega)$' which is technically nonsense in the context of the second definition. Let the reader be warned that we will generally commit this sin. Thus when we are talking about L^p-functions and we write $f = g$ we really have in mind that f and g are two functions that agree μ-almost everywhere. If the context is changed to, say, continuous functions, then $f = g$ means $f(x) = g(x)$ for all x. In particular, we note that it makes no sense to ask for the value $f(0)$, say, if f is an L^p-function.

● A **convex set** $K \subset \mathbb{R}^n$ is one for which $\lambda x+(1-\lambda)y \in K$ for all $x, y \in K$ and all $0 \le \lambda \le 1$. A **convex function**, f, on a convex set $K \subset \mathbb{R}^n$ is a real-valued function satisfying

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \tag{8}$$

for all $x, y \in K$ and all $0 \le \lambda \le 1$. If equality *never* holds in (8) when $y \neq x$ and $0 < \lambda < 1$, then f is **strictly convex**. More generally, we say that f is strictly convex at a point $x \in K$ if $f(x) < \lambda f(y) + (1-\lambda) f(z)$ whenever $x = \lambda y + (1-\lambda)z$ for $0 < \lambda < 1$ and $y \neq z$. If the inequality (8) is reversed, f is said to be **concave** (alternatively, f is concave $\iff$ $-f$ is convex). It is easy to prove that if K is an *open* set, then a convex function is continuous.

A **support plane** to a graph of a function $f : K \to \mathbb{R}$ at a point $x \in K$ is a plane (in $\mathbb{R}^{n+1}$) that touches the graph at $(x, f(x))$ and that nowhere lies above the graph. In general, a support plane might not exist at x, but if f is convex on K, its graph has at least one support plane at each point of the interior of K. Thus there exists a vector $V \in \mathbb{R}^n$ (which depends on x) such that

$$f(y) \ge f(x) + V \cdot (y - x) \tag{9}$$

for all $y \in K$. If the support plane at x is unique it is called a **tangent plane**. If f is convex, the existence of a tangent plane at x is equivalent to differentiability at x.

If $n = 1$ and if f is convex, f need not be differentiable at x. However, when x is in the interior of the interval K, f always has a **right derivative**, $f'_+(x)$, and a **left derivative**, $f'_-(x)$, at x, e.g.,

$$f'_+(x) := \lim_{\varepsilon \searrow 0} [f(x+\varepsilon) - f(x)]/\varepsilon.$$

See [Hardy–Littlewood–Pólya] and Exercise 18.

2.2 THEOREM (Jensen's inequality)

Let $J : \mathbb{R} \to \mathbb{R}$ be a convex function. Let f be a real-valued function on some set Ω that is measurable with respect to some Σ-algebra, and let μ be a measure on Σ. Since J is convex, it is continuous and therefore $(J \circ f)(x) := J(f(x))$ is also a Σ-measurable function on Ω. We assume that $\mu(\Omega) = \int_\Omega \mu(\mathrm{d}x)$ is finite.

Suppose now that $f \in L^1(\Omega)$ and let $\langle f \rangle$ be the average of f, i.e.,

$$\langle f \rangle = \frac{1}{\mu(\Omega)} \int_\Omega f \, \mathrm{d}\mu.$$

Then

(i) $[J \circ f]_-$, *the negative part of* $[J \circ f]$, *is in* $L^1(\Omega)$, *whence* $\int_\Omega (J \circ f)(x)\mu(\mathrm{d}x)$ *is well defined although it might be* $+\infty$.

(ii)
$$\langle J \circ f \rangle \geq J(\langle f \rangle). \tag{1}$$

If J is strictly convex at $\langle f \rangle$ there is equality in (1) *if and only if f is a constant function.*

PROOF. Since J is convex its graph has at least one support line at each point. Thus, there is a constant $V \in \mathbb{R}$ such that

$$J(t) \geq J(\langle f \rangle) + V(t - \langle f \rangle) \tag{2}$$

for all $t \in \mathbb{R}$. From this we conclude that

$$[J(f)]_-(x) \leq |J(\langle f \rangle)| + |V||\langle f \rangle| + |V||f(x)|,$$

and hence, recalling that $\mu(\Omega) < \infty$, (i) is proved.

If we now substitute $f(x)$ for t in (2) and integrate over Ω we arrive at (1).

Assume now that J is strictly convex at $\langle f \rangle$. Then (2) is a strict inequality either for all $t > \langle f \rangle$ or for all $t < \langle f \rangle$. If f is not a constant, then $f(x) - \langle f \rangle$ takes on both positive and negative values on sets of positive measure. This implies the last assertion of the theorem. ■

● The importance of the next inequality can hardly be overrated. There are many proofs of it and the one we give is not necessarily the simplest; we give it in order to show how the inequality is related to Jensen's inequality. Another proof is outlined in the exercises.

2.3 THEOREM (Hölder's inequality)

Let p and q be dual indices, i.e., $1/p + 1/q = 1$ with $1 \leq p \leq \infty$. Let $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$. Then the pointwise product, given by $(fg)(x) = f(x)g(x)$, is in $L^1(\Omega)$ and

$$\left| \int_\Omega fg \, \mathrm{d}\mu \right| \leq \int_\Omega |f||g| \, \mathrm{d}\mu \leq \|f\|_p \|g\|_q. \tag{1}$$

The first inequality in (1) *is an equality if and only if*

(i) $f(x)g(x) = e^{i\theta}|f(x)||g(x)|$ *for some real constant* θ *and for* μ-*almost every* x.

If $f \not\equiv 0$ *the second inequality in* (1) *is an equality if and only if there is a constant* $\lambda \in \mathbb{R}$ *such that:*

(iia) *If* $1 < p < \infty$, $|g(x)| = \lambda |f(x)|^{p-1}$ *for* μ-*almost every* x.

(iib) *If* $p = 1$, $|g(x)| \leq \lambda$ *for* μ-*almost every* x *and* $|g(x)| = \lambda$ *when* $f(x) \neq 0$.

(iic) *If* $p = \infty$, $|f(x)| \leq \lambda$ *for* μ-*almost every* x *and* $|f(x)| = \lambda$ *when* $g(x) \neq 0$.

REMARKS. (1) The special case $p = q = 2$ is the **Schwarz inequality**

$$\left| \int_\Omega fg \right|^2 \leq \int_\Omega |f|^2 \int_\Omega |g|^2. \tag{2}$$

(2) If $f_1, \ldots, f_m$ are functions on Ω with $f_i \in L^{p_i}(\Omega)$ and $\sum_{j=1}^m 1/p_i = 1$ then

$$\left| \int_\Omega \prod_{j=1}^m f_i \, \mathrm{d}\mu \right| \leq \prod_{j=1}^m \|f_i\|_{p_i}. \tag{3}$$

This generalization is a simple consequence of (1) with $f := f_1$ and $g := \prod_{j=2}^m f_j$. Then use induction on $\int_\Omega |g|^p$.

PROOF. The left inequality in (1) is a triviality, so we may as well suppose $f \geq 0$ and $g \geq 0$ (note that condition (i) is what is needed for equality here). The cases $p = \infty$ and $q = \infty$ are trivial so we suppose that $1 < p, q < \infty$. Set $A = \{x : g(x) > 0\} \subset \Omega$ and let $B = \Omega \sim A = \{x : g(x) = 0\}$. Since

$$\int_\Omega f^p \, \mathrm{d}\mu = \int_A f^p \, \mathrm{d}\mu + \int_B f^p \, \mathrm{d}\mu,$$

since $\int_\Omega g^p \, \mathrm{d}\mu = \int_A g^p \, \mathrm{d}\mu$, and since $\int_\Omega fg \, \mathrm{d}\mu = \int_A fg \, \mathrm{d}\mu$, we see that it suffices—in order to prove (1)—to assume that $\Omega = A$. (Why is $\int fg \, \mathrm{d}\mu$ defined?) Introduce a new measure on $\Omega = A$ by $\nu(\mathrm{d}x) = g(x)^q \mu(\mathrm{d}x)$. Also, set $F(x) = f(x)g(x)^{-q/p}$ (which makes sense since $g(x) > 0$ a.e.). Then, with respect to the measure ν, we have that $\langle F \rangle = \int_\Omega fg \, \mathrm{d}\mu / \int_\Omega g^q \, \mathrm{d}\mu$. On the other hand, with $J(t) = |t|^p$, $\int_\Omega J \circ F \, \mathrm{d}\nu = \int_\Omega f^p \, \mathrm{d}\mu$. Our conclusion (1) is then an immediate consequence of Jensen's inequality—as is the condition for equality. ■

2.4 THEOREM (Minkowski's inequality)

Suppose that Ω and Γ are any two spaces with sigma-finite measures μ and ν respectively. Let f be a nonnegative function on $\Omega \times \Gamma$ which is $\mu \times \nu$-measurable. Let $1 \leq p < \infty$. Then

$$\begin{aligned} &\int_\Gamma \left(\int_\Omega f(x,y)^p \mu(\mathrm{d}x) \right)^{1/p} \nu(\mathrm{d}y) \\ &\qquad \geq \left(\int_\Omega \left(\int_\Gamma f(x,y)\nu(\mathrm{d}y) \right)^p \mu(\mathrm{d}x) \right)^{1/p} \end{aligned} \tag{1}$$

in the sense that the finiteness of the left side implies the finiteness of the right side.

Equality and finiteness in (1) for $1 < p < \infty$ imply the existence of a μ-measurable function $\alpha : \Omega \to \mathbb{R}^+$ and a ν-measurable function $\beta : \Gamma \to \mathbb{R}^+$ such that

$$f(x,y) = \alpha(x)\beta(y) \quad \textit{for } \mu \times \nu\textit{-almost every } (x,y).$$

A special case of this is the **triangle inequality**. *For $f, g \in L^p(\Omega, \mathrm{d}\mu)$ (possibly complex functions)*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p \quad \textit{for } 1 \leq p \leq \infty. \tag{2}$$

If $f \not\equiv 0$ and if $1 < p < \infty$, there is equality in (2) if and only if $g = \lambda f$ for some $\lambda \geq 0$.

PROOF. First we note that the two functions

$$\int_\Omega f(x,y)^p \mu(\mathrm{d}x) \quad \text{and} \quad H(x) := \int_\Gamma f(x,y)\nu(\mathrm{d}y)$$

are measurable functions. This follows from Theorem 1.12 (Fubini's theorem) and the assumption that f is $\mu \times \nu$-measurable. We can assume that $f > 0$ on a set of positive $\mu \times \nu$ measure, for otherwise there is nothing to prove. We can also assume that the right side of (1) is finite; if not we can truncate f so that it is finite and then use a monotone convergence argument to remove the truncation. Sigma-finiteness is again used in this step.

The right side of (1) can be written as follows:

$$\int_\Omega H(x)^p \mu(\mathrm{d}x) = \int_\Omega \left(\int_\Gamma f(x,y)\nu(\mathrm{d}y) \right) H(x)^{p-1}\mu(\mathrm{d}x)$$
$$= \int_\Gamma \left(\int_\Omega f(x,y) H(x)^{p-1}\mu(\mathrm{d}x) \right) \nu(\mathrm{d}y).$$

The last equation follows by Fubini's theorem. Using Theorem 2.3 (Hölder's inequality) on the right side we obtain

$$\int_\Omega H(x)^p \mu(\mathrm{d}x) \le \int_\Gamma \left(\int_\Omega f(x,y)^p \mu(\mathrm{d}x) \right)^{1/p}$$
$$\times \left(\int_\Omega H(x)^p \mu(\mathrm{d}x) \right)^{\frac{p-1}{p}} \nu(\mathrm{d}y). \tag{3}$$

Dividing both sides of (3) by

$$\left(\int_\Omega H(x)^p \mu(\mathrm{d}x) \right)^{(p-1)/p},$$

which is neither zero nor infinity (by our assumptions about f), yields (1).

The equality sign in the use of Hölder's inequality implies that for ν-almost every y there exists a number $\lambda(y)$ (i.e., independent of x) such that

$$\lambda(y) H(x) = f(x,y) \text{ for } \mu\text{-almost every } x. \tag{4}$$

As mentioned above, H is μ-measurable. To see that λ is ν-measurable we note that

$$\lambda(y) \int_\Omega H(x)^p \mu(\mathrm{d}x) = \int_\Omega f(x,y)^p \mu(\mathrm{d}x),$$

and this yields the desired result since the right side is ν-measurable (by Fubini's theorem).

It remains to prove (2). First, by observing that

$$|f(x) + g(x)| \le |f(x)| + |g(x)|, \tag{5}$$

the problem is reduced to proving (2) for nonnegative functions. Evidently, (5) implies (2) when $p = 1$ or ∞, so we can assume $1 < p < \infty$. We set $F(x,1) = |f(x)|$, $F(x,2) = |g(x)|$ and let ν be the counting measure of the set $\Gamma = \{1,2\}$, namely $\nu(\{1\}) = \nu(\{2\}) = 1$. Then the inequality (2) is seen to be a special case of (1). (Note the use of Fubini's theorem here.)

Equality in (2) entails the existence of constants λ_1 and λ_2 (independent of x) such that

$$|f(x)| = \lambda_1(|f(x)| + |g(x)|) \quad \text{and} \quad |g(x)| = \lambda_2(|f(x)| + |g(x)|). \tag{6}$$

Thus, $|g(x)| = \lambda |f(x)|$ almost everywhere for some constant λ. However, equality in (5) means that $g(x) = \lambda f(x)$ with λ real and nonnegative. ■

● If $1 < p < \infty$, then $L^p(\Omega)$ possesses another geometric structure that has many consequences, among them the characterization of the dual of $L^p(\Omega)$ (2.14) and, in connection with weak convergence, Mazur's theorem (2.13). This structure is called **uniform convexity** and will be described next. The version we give is optimal and is due to [Hanner]; the proof is in [Ball-Carlen-Lieb]. It improves the triangle (or convexity) inequality

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

2.5 THEOREM (Hanner's inequality)

Let f and g be functions in $L^p(\Omega)$. If $1 \le p \le 2$, we have

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

$$(\|f+g\|_p + \|f-g\|_p)^p + \big|\|f+g\|_p - \|f-g\|_p\big|^p \le 2^p(\|f\|_p^p + \|g\|_p^p). \tag{2}$$

If $2 \le p < \infty$, the inequalities are reversed.

REMARK. When $\|f\|_p = \|g\|_p$, (2) improves the triangle inequality $\|f+g\|_p \le \|f\|_p + \|g\|_p$ because, by convexity of $t \mapsto |t|^p$, the left side of (2) is not smaller than $2\|f+g\|_p^p$. To be more precise, it is easy to prove (Exercise 4) that the left side of (2) is bounded below for $1 \le p \le 2$ and for $\|f-g\|_p \le \|f+g\|_p$ by

$$2\|f+g\|_p^p + p(p-1)\|f+g\|_p^{p-2}\|f-g\|_p^2.$$

The geometric meaning of Theorem 2.5 is explored in Exercise 5.

PROOF. (1) and (2) are identities when $p = 2$ ((1) is then called the **parallelogram identity**) and reduce to the triangle inequality if $p = 1$. (2) is derived from (1) by the replacements $f \to f+g$ and $g \to f-g$. Thus, we concentrate on proving (1) for $p \ne 2$. We can obviously assume that $R := \|g\|_p/\|f\|_p \le 1$ and that $\|f\|_p = 1$. For $0 \le r \le 1$ define

$$\alpha(r) = (1+r)^{p-1} + (1-r)^{p-1}$$

and

$$\beta(r) = [(1+r)^{p-1} - (1-r)^{p-1}]r^{1-p},$$

with $\beta(0) = 0$ for $p < 2$ and $\beta(0) = \infty$ for $p > 2$. We first claim that the function $F_R(r) = \alpha(r) + \beta(r)R^p$ has its maximum at $r = R$ (if $p < 2$) and its minimum at $r = R$ (if $p > 2$). In both cases $F_R(R) = (1+R)^p + (1-R)^p$. To prove this assertion we can use the calculus to compute

$$\begin{aligned} dF_R(r)/dr &= \alpha'(r) + \beta'(r)R^p \\ &= (p-1)[(1+r)^{p-2} - (1-r)^{p-2}](1 - (R/r)^p), \end{aligned}$$

which shows that the derivative of $F_R(r)$ vanishes only at $r = R$ and that the sign of the derivative for $r \neq R$ is such that the point $r = R$ is a maximum or minimum as stated above. Furthermore, for all $0 \le r \le 1$ we have that $\beta(r) \le \alpha(r)$ (if $p < 2$) and $\beta(r) \ge \alpha(r)$ (if $p > 2$) and thus, when $R > 1$,

$$\alpha(r) + \beta(r)R^p \le \alpha(r)R^p + \beta(r) \ (\text{if } p < 2)$$

and

$$\alpha(r) + \beta(r)R^p \ge \alpha(r)R^p + \beta(r) \ (\text{if } p > 2).$$

Thus, in all cases we have for all $0 \le r \le 1$ and all nonnegative numbers A and B

$$\alpha(r)|A|^p + \beta(r)|B|^p \le |A+B|^p + |A-B|^p, \qquad p < 2, \tag{3}$$

and the reverse if $p > 2$. It is important to note that equality holds if $r = B/A \le 1$.

In fact, (3) and its reverse for $p > 2$ hold for complex A and B (that is why we wrote (3) with $|A|, |B|$, etc.). To see this note that it suffices to prove it when $A = a$ and $B = be^{i\theta}$ with $a, b > 0$. It then suffices to show that $(a^2 + b^2 + 2ab\cos\theta)^{p/2} + (a^2 + b^2 - 2ab\cos\theta)^{p/2}$ has its minimum when $\theta = 0$ (if $p < 2$) or its maximum when $\theta = 0$ (if $p > 2$). But this follows from the fact that the function $x \mapsto x^r$ is concave (if $0 < r < 1$) or convex (if $r > 1$).

To prove (1) it suffices, then, to prove that when $1 \le p < 2$

$$\int \{|f+g|^p + |f-g|^p\} \, \mathrm{d}\mu \ge \alpha(r) \int |f|^p \, \mathrm{d}\mu + \beta(r) \int |g|^p \, \mathrm{d}\mu \tag{4}$$

for every $0 \le r \le 1$, and the reverse inequality when $p > 2$. But to prove (4) it suffices to prove it pointwise, i.e., for complex numbers f and g. That is, we have to prove

$$|f+g|^p + |f-g|^p \ge \alpha(r)|f|^p + \beta(r)|g|^p \quad \text{for } p < 2$$

(and the reverse for $p > 2$). But this follows from (3). ∎

● Differentiability of $\|f+tg\|_p^p = \int |f+tg|^p$ with respect to $t \in \mathbb{R}$ will prove to be useful. Note that this function of t is convex and hence always has a left and right derivative. In case $p = 1$ it may not be truly differentiable, however, but it is so for $p > 1$, as we show next.

2.6 THEOREM (Differentiability of norms)

Suppose f and g are functions in $L^p(\Omega)$ with $1 < p < \infty$. The function defined on $\mathbb{R}$ by

$$N(t) = \int_\Omega |f(x) + tg(x)|^p \mu(\mathrm{d}x)$$

is differentiable and its derivative at $t = 0$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} N\bigg|_{t=0} = \frac{p}{2} \int_\Omega |f(x)|^{p-2} \{\overline{f}(x) g(x) + f(x) \overline{g}(x)\} \mu(\mathrm{d}x). \tag{1}$$

REMARKS. (1) Note that $|f|^{p-2} f$ is well defined for $1 < p$, even when $f = 0$, in which case it equals 0. This convention will occur frequently in the sequel. Note also that $|f|^{p-2} f$ and $|f|^{p-2} \overline{f}$ are functions in $L^{p'}(\Omega)$.

(2) This notion of derivative of the norm is called the **Gateaux- or directional derivative**.

PROOF. It is an elementary fact from calculus that for complex numbers f and g we have

$$\lim_{t \to 0} [|f + tg|^p - |f|^p]/t = \frac{p}{2} |f|^{p-2} (\overline{f} g + f \overline{g}),$$

i.e., $|f + tg|^p$ is differentiable. Our problem, then, is to interchange differentiation and integration. To do so we use the inequality (for $|t| \leq 1$)

$$|f|^p - |f - g|^p \leq \frac{1}{t} \{|f + tg|^p - |f|^p\} \leq |f + g|^p - |f|^p,$$

which follows from the convexity of $x \to x^p$ (e.g., $|f + tg|^p \leq (1 - t)|f|^p + t|f + g|^p$). Since $|f|^p, |f + g|^p$ and $|f - g|^p$ are fixed, summable functions, we can do the necessary interchange thanks to the dominated convergence theorem. ■

2.7 THEOREM (Completeness of L^p-spaces)

Let $1 \le p \le \infty$ and let f^i, for $i = 1,2,3,\ldots$, be a **Cauchy sequence** *in $L^p(\Omega)$, i.e., $\|f^i - f^j\|_p \to 0$ as $i,j \to \infty$. (This means that for each $\varepsilon > 0$ there is an N such that $\|f^i - f^j\|_p < \varepsilon$ when $i > N$ and $j > N$.) Then there exists a unique function $f \in L^p(\Omega)$ such that $\|f^i - f\|_p \to 0$ as $i \to \infty$. We denote this latter fact by*

$$f^i \to f \quad \textit{as} \quad i \to \infty,$$

and we say that f^i **converges strongly to f in $L^p(\Omega)$.**

Moreover, there exists a subsequence $f^{i_1}, f^{i_2}, \ldots$ (with $i_1 < i_2 < \cdots$, of course) and a nonnegative function F in $L^p(\Omega)$ such that

(i) *Domination :* $|f^{i_k}(x)| \le F(x)$ *for all k and μ-almost every x.* (1)

(ii) *Pointwise convergence:* $\lim\limits_{k\to\infty} f^{i_k}(x) = f(x)$ *for μ-almost every x.* (2)

REMARK. 'Convergence' and 'strong convergence' are used interchangeably. The phrase norm convergence is also used.

PROOF. The first, and most important remark, concerns a strategy that is frequently very useful. Namely, it suffices to show the strong convergence for *some* subsequence. To prove this sufficiency, let f^{i_k} be a subsequence that converges strongly to f in $L^p(\Omega)$ as $k \to \infty$. Since, by the triangle inequality,

$$\|f^i - f\|_p \le \|f^i - f^{i_k}\|_p + \|f^{i_k} - f\|_p,$$

we see that for any $\varepsilon > 0$ we can make the last term on the right side less than $\varepsilon/2$ by choosing k large. The first term on the right can be made smaller than $\varepsilon/2$ by choosing i and k large enough, since f^i is a Cauchy sequence. Thus, $\|f^i - f\|_p < \varepsilon$ for i large enough and we can conclude convergence for the *whole sequence,* i.e., $f^i \to f$. This also proves, incidentally, that the limit—if it exists—is unique.

To obtain such a convergent subsequence pick a number i_1 such that $\|f^{i_1} - f^n\|_p \le 1/2$ for all $n \ge i_1$. That this is possible is precisely the definition of a Cauchy sequence. Now choose i_2 such that $\|f^{i_2} - f^n\|_p < 1/4$ for all $n \ge i_2$ and so on. Thus we have obtained a subsequence of the integers, i_k, with the property that $\|f^{i_k} - f^{i_{k+1}}\|_p \le 2^{-k}$ for $k = 1,2,\ldots$. Consider the monotone sequence of positive functions

$$F_l(x) := |f^{i_1}(x)| + \sum_{k=1}^{l} |f^{i_k}(x) - f^{i_{k+1}}(x)|. \tag{3}$$

By the triangle inequality

$$\|F_l\|_p \le \|f^{i_1}\|_p + \sum_{k=1}^{l} 2^{-k} \le \|f^{i_1}\|_p + 1.$$

Thus, by the monotone convergence theorem, F_l converges pointwise μ-a.e. to a positive function F which is in $L^p(\Omega)$ and hence is finite almost everywhere. The sequence

$$f^{i_{k+1}}(x) = f^{i_1}(x) + (f^{i_2}(x) - f^{i_1}(x)) + \cdots + (f^{i_{k+1}}(x) - f^{i_k}(x)) \qquad (4)$$

thus converges absolutely for almost every x, and hence it also converges for the same x's to some number $f(x)$. Since $|f^{i_k}(x)| \le F(x)$ and $F \in L^p(\Omega)$, we know by dominated convergence that f is in $L^p(\Omega)$. Again by dominated convergence $\|f^{i_k} - f\|_p \to 0$ as $k \to \infty$ since $|f^{i_k}(x) - f(x)| \le F(x) + |f(x)| \in L^p(\Omega)$. Thus, the subsequence f^{i_k} converges strongly in $L^p(\Omega)$ to f. ■

● An example of the use of uniform convexity, Theorem 2.5, is provided by the following projection lemma, which will be useful later.

2.8 LEMMA (Projection on convex sets)

Let $1 < p < \infty$ and let K be a convex set in $L^p(\Omega)$ (i.e., $f, g \in K \Rightarrow tf + (1-t)g \in K$ for all $0 \le t \le 1$) which is also a **norm closed set** *(i.e., if $\{g^i\}$ is a Cauchy sequence in K, then its limit, g, is also in K). Let $f \in L^p(\Omega)$ be any function that is not in K and define the distance as*

$$D = \ \mathrm{dist}(f, K) = \inf_{g \in K} \|f - g\|_p. \qquad (1)$$

Then there is a function $h \in K$ such that $D = \|f - h\|_p$.

Every function $g \in K$ satisfies

$$\mathrm{Re} \int_\Omega [(g - h)(\overline{f} - \overline{h})]|f - h|^{p-2}\, \mathrm{d}\mu \le 0. \qquad (2)$$

PROOF. We shall prove this for $p \le 2$ using the uniform convexity result 2.5(2) and shall assume $f = 0$. We leave the rest to the reader. Let $h^j, j = 1, 2, \ldots$ be a minimizing sequence in K, i.e., $\|h^j\|_p \to D$. We shall show that this is a Cauchy sequence. First note that $\|h^j + h^k\|_p \to 2D$ as $j, k \to \infty$ (because $\|h^j + h^k\|_p \le \|h^j\|_p + \|h^k\|_p$, which converges to $2D$, but $\|h^j + h^k\|_p \ge 2D$ since $\frac{1}{2}(h^j + h^k) \in K$). From 2.5(2) we have that

$$(\|h^j + h^k\|_p + \|h^j - h^k\|_p)^p + \big|\|h^j + h^k\|_p - \|h^j - h^k\|_p\big|^p \le 2^p\{\|h^j\|_p^p + \|h^k\|_p^p\}.$$

The right side converges as $j, k \to \infty$ to $2^{p+1}D^p$. Suppose that $\|h^j - h^k\|_p$ does not tend to zero, but instead (for infinitely many j's and k's) stays bounded below by some number $b > 0$. Then we would have

$$|2D + b|^p + |2D - b|^p \le 2^{p+1}D^p,$$

which implies that $b = 0$ (by the *strict* convexity of $x \to |2D + x|^p$, which implies that $|2D+x|^p+|2D-x|^p > 2|2D|^p$ unless $x = 0$). Thus, our sequence is Cauchy and, since K is closed, it has a limit $h \in K$.

To verify (2) we fix $g \in K$ and set $g_t = (1-t)h + tg \in K$ for $0 \le t \le 1$. Then (with $f = 0$ as before) $N(t) := \|f - g_t\|_p^p \ge D^p$ while $N(0) = D^p$. Since $N(t)$ is differentiable (Theorem 2.6) we have that $N'(0) \ge 0$, and this is exactly (2) (using 2.6(1)). ■

2.9 DEFINITION (Continuous linear functionals and weak convergence)

The notion of strong convergence just mentioned in Theorem 2.7 (completeness of L^p-spaces) is not the only useful notion of convergence in $L^p(\Omega)$. The second notion, **weak convergence**, requires continuous linear functionals—which we now define. (Incidentally, what is said here applies to any normed vector space—not just $L^p(\Omega)$.) Weak convergence is often more useful than strong convergence for the following reason. We know that a closed, bounded set, A, in $\mathbb{R}^n$ is compact, i.e., every sequence $x^1, x^2, \dots$ in A has a subsequence with a limit in A. *The analogous compactness assertion in $L^p(\mathbb{R}^n)$, or even $L^p(\Omega)$ for Ω a compact set in $\mathbb{R}^n$, is false.* Below, we show how to construct a sequence of functions, bounded in $L^p(\mathbb{R}^n)$ for every p, but for which there is no convergent subsequence in *any* $L^p(\mathbb{R}^n)$.

If weak convergence is substituted for strong convergence, the situation improves. *The main theorem here, toward which we are headed, is the Banach–Alaoglu Theorem* 2.18 *which shows that the bounded sets are compact, with this notion of weak convergence, when* $1 < p < \infty$.

A map, L, from $L^p(\Omega)$ to the complex numbers is a **linear functional** if

$$L(af_1 + bf_2) = aL(f_1) + bL(f_2) \tag{1}$$

for all $f_1, f_2 \in L^p(\Omega)$ and $a, b \in \mathbb{C}$. It is a **continuous linear functional** if, for every strongly convergent sequence, f^i,

$$L(f^i) \to L(f) \quad \text{when} \quad f^i \to f. \tag{2}$$

It is a **bounded linear functional** if

$$|L(f)| \le K\|f\|_p \tag{3}$$

for some finite number K. We leave it as a very easy exercise for the reader to prove that

$$\text{bounded} \quad \Longleftrightarrow \quad \text{continuous} \tag{4}$$

for linear maps.

The set of *continuous* linear functionals (continuity is crucial) on $L^p(\Omega)$ is called the **dual** of $L^p(\Omega)$ and is denoted by $L^p(\Omega)^*$. It is also a vector space over the complex numbers (since sums and scalar multiples of elements of $L^p(\Omega)^*$ are in $L^p(\Omega)^*$). This new space has a norm defined by

$$\|L\| = \sup\{|L(f)| : \|f\|_p \le 1\}. \tag{5}$$

The reader is asked to check that this definition (5) has the three crucial properties of a norm given in 2.1 (a,b,c) : $\|\lambda L\| = |\lambda|\|L\|$, $\|L\| = 0 \Leftrightarrow L = 0$, and the triangle inequality.

It is important to know all the elements of the dual of $L^p(\Omega)$ (or any other vector space). The reason is that an element $f \in L^p(\Omega)$ can be uniquely identified (as we shall see in Theorem 2.10 (linear functionals separate)) if we know how *all* the elements of the dual act on f, i.e., if we know $L(f)$ for all $L \in L^p(\Omega)^*$.

Weak convergence.

If $f, f^1, f^2, f^3, \ldots$ is a sequence of functions in $L^p(\Omega)$, we say that f^i **converges weakly to** f (and write $f^i \rightharpoonup f$) if

$$\lim_{i\to\infty} L(f^i) = L(f) \tag{6}$$

for *every* $L \in L^p(\Omega)^*$.

An obvious but important remark is that strong convergence implies weak convergence, i.e., if $\|f^i - f\|_p \to 0$ as $i \to \infty$, then $\lim_{i\to\infty} L(f^i) = L(f)$ for all continuous linear functionals L. In particular, strong limits and weak limits have to agree, if they both exist (cf. Theorem 2.10).

Two questions that immediately present themselves are (a) what is $L^p(\Omega)^*$ and (b) how is it possible for f^i to converge weakly, but not strongly, to f? For the former, Hölder's inequality (Theorem 2.3) immediately implies that $L^{p'}(\Omega)$ is a subset of $L^p(\Omega)^*$ when $\frac{1}{p'} + \frac{1}{p} = 1$. A function $g \in L^{p'}(\Omega)$ acts on arbitrary functions $f \in L^p(\Omega)$ by

$$L_g(f) = \int_\Omega g(x) f(x) \mu(\mathrm{d}x). \tag{7}$$

It is easy to check that L_g is linear and continuous. A deeper question is whether (7) gives us *all of* $L^p(\Omega)^*$. The answer will turn out to be 'yes' for $1 \le p < \infty$, and 'no' for $p = \infty$.

If we accept this conclusion for the moment we can answer question (b) above in the following heuristic way when $\Omega = \mathbb{R}^n$ and $1 < p < \infty$. There are three basic mechanisms by which $f^k \rightharpoonup f$ but $f^k \not\to f$ and we illustrate each for $n = 1$.

(i) f^k 'oscillates to death': An example is $f^k(x) = \sin kx$ for $0 \leq x \leq 1$ and zero otherwise.

(ii) f^k 'goes up the spout': An example is $f^k(x) = k^{1/p} g(kx)$, where g is any fixed function in $L^p(\mathbb{R}^1)$. This sequence becomes very large near $x = 0$.

(iii) f^k 'wanders off to infinity': An example is $f^k(x) = g(x + k)$ for some fixed function g in $L^p(\mathbb{R}^1)$.

In each case $f^k \rightharpoonup 0$ weakly but f^k does not converge strongly to zero (or to anything else). We leave it to the reader to prove this assertion; some of the theorems proved later in this section will be helpful.

We begin our study of weak convergence by showing that there are enough elements of $L^p(\Omega)^*$ to identify all elements of $L^p(\Omega)$. Much of what we prove here is normally proved with the **Hahn–Banach theorem**. We do not use it for several reasons. One is that the interested reader can easily find it in many texts. Another reason is that it is not necessary in the case of $L^p(\Omega)$ spaces and we prefer a direct 'hands on' approach to an abstract approach—wherever the abstract approach does not add significant enlightenment.

2.10 THEOREM (Linear functionals separate)

Suppose that $f \in L^p(\Omega)$ satisfies

$$L(f) = 0 \quad \textit{for all } L \in L^p(\Omega)^*. \tag{1}$$

(In the case $p = \infty$ we also assume that our measure space is sigma-finite, but this restriction can be lifted by invoking transfinite induction.) Then

$$f = 0.$$

Consequently, if $f^i \rightharpoonup k$ and $f^i \rightharpoonup h$ weakly in $L^p(\Omega)$, then $k = h$.

PROOF. If $1 < p < \infty$ define

$$g(x) = |f(x)|^{p-2} \overline{f}(x)$$

when $f(x) \neq 0$, and set $g(x) = 0$ otherwise. The fact that $f \in L^p(\Omega)$ immediately implies that $g \in L^{p'}(\Omega)$. We also have that $\int gf = \|f\|_p^p$.

But, as we said in 2.9(7), the functional $h \to \int gh$ is a continuous linear functional. Hence, $\int gf = \|f\|_p = 0$ by our hypothesis (1), which implies $f = 0$.

If $p = 1$ we take

$$g(x) = \overline{f}(x)/|f(x)|$$

if $f(x) \neq 0$, and $g(x) = 0$ otherwise. Then $g \in L^\infty(\Omega)$ and the above argument applies. If $p = \infty$ set $A = \{x : |f(x)| > 0\}$. If $f \not\equiv 0$, then $\mu(A) > 0$. Take any measurable subset $B \subset A$ such that $0 < \mu(B) < \infty$; such a set exists by sigma-finiteness. Set $g(x) = \overline{f}(x)/|f(x)|$ for $x \in B$ and zero otherwise. Clearly, $g \in L^1(\Omega)$ and the previous argument can be applied. ■

2.11 THEOREM (Lower semicontinuity of norms)

For $1 \leq p \leq \infty$ the L^p-norm is **weakly lower semicontinuous**, *i.e., if $f^j \rightharpoonup f$ weakly in $L^p(\Omega)$, then*

$$\liminf_{j\to\infty} \|f^j\|_p \geq \|f\|_p. \tag{1}$$

If $p = \infty$ we make the extra technical assumption that the measure μ is sigma finite.

Moreover, if $1 < p < \infty$ and if $\lim_{j\to\infty} \|f^j\|_p = \|f\|_p$, then $f^j \to f$ strongly as $j \to \infty$.

REMARK. The second part of this theorem is very useful in practice because it often provides a way to identify strongly convergent sequences. For the connection with semicontinuity as in Sect. 1.5, cf. Exercise 1.2. Compare, also, Remark (2) after Theorem 1.9.

PROOF. For $1 \leq p < \infty$ consider the functional

$$L(h) = \int gh \quad \text{with} \quad g(x) = |f(x)|^{p-2}\overline{f}(x)$$

as in the proof of the separation theorem, Theorem 2.10. Since $L(f) = \|f\|_p^p$, we have, by Hölder's inequality with $1/p + 1/q = 1$,

$$\|f\|_p^p = \lim_{j\to\infty} L(f^j) \leq \|g\|_q \liminf_{j\to\infty} \|f^j\|_p$$

which, since $\|g\|_q = \|f\|_p^{p-1}$, gives (1).

For $p = \infty$ assume $\|f\|_\infty =: a > 0$ and consider the set

$$A_\varepsilon = \{x \in \Omega : |f(x)| > a - \varepsilon\}.$$

Since the space (Ω, μ) is sigma-finite, there is a sequence of sets B_k of finite measure such that $A_\varepsilon \cap B_k$ increases to A_ε. Set $g_{k,\varepsilon} = f(x)/|f(x)|$ if $x \in A_\varepsilon \cap B_k$ and zero otherwise. Now by Hölder's inequality

$$\mu(A_\varepsilon \cap B_k) \liminf_{j\to\infty} \|f^j\|_\infty \ \geq \ \lim_{j\to\infty} \int g_{k,\varepsilon} f^j = \int_{A_\varepsilon \cap B_k} |f(x)| \, \mathrm{d}\mu,$$

where the last equation follows from the weak convergence of f^j to f. But

$$\int_{A_\varepsilon \cap B_k} |f(x)| \, \mathrm{d}\mu \geq (a - \varepsilon)\mu(A_\varepsilon \cap B_k),$$

and hence $\liminf_{j\to\infty} \|f^j\|_\infty \geq \|f\|_\infty - \varepsilon$ for all $\varepsilon > 0$.

Thus far we have proved (1). To prove the second assertion for $1 < p < \infty$ we first note that $\lim \|f^j\|_p = \|f\|_p$ implies that $\lim \|f^j + f\|_p = 2\|f\|_p$ (clearly $f^j + f \rightharpoonup 2f$ and, by (1), $\liminf \|f^j + f\|_p \geq 2\|f\|_p$, but $\|f^j + f\|_p \leq \|f^j\|_p + \|f\|_p$ by the triangle inequality). For $p \leq 2$ we use the uniform convexity 2.5(2) (we leave $p > 2$ to the reader) with $g = f^j$. Taking limits we have (with $A_j = \|f + f^j\|_p$ and $B_j = \|f - f^j\|_p$)

$$\limsup_{j\to\infty} \left\{(A_j + B_j)^p + |A_j - B_j|^p\right\} \leq 2^{p+1}\|f\|_p^p.$$

Since $x \mapsto |A + x|^p$ is *strictly* convex for $1 < p < \infty$, and since $A_j \to 2\|f\|_p$, B_j must tend to zero. ∎

● The next theorem shows that weakly convergent sequences are, at least, norm bounded.

2.12 THEOREM (Uniform boundedness principle)

Let $f^1, f^2, \ldots$ be a sequence in $L^p(\Omega)$ with the following property: For each functional $L \in L^p(\Omega)^*$ the sequence of numbers $L(f^1), L(f^2), \ldots$ is bounded. Then the norms $\|f^j\|_p$ are bounded, i.e., $\|f^j\|_p < C$ for some finite $C > 0$.

PROOF. We suppose the theorem is false and will derive a contradiction. We do this for $1 < p < \infty$, and leave the easy modifications for $p = 1$ and $p = \infty$ to the reader.

First, for the following reason, we can assume that $\|f^j\|_p = 4^j$. By choosing a subsequence (which we continue to denote by $j = 1, 2, 3, \dots$) we can certainly arrange that $\|f^j\|_p \geq 4^j$. Then we replace the sequence f^j by the sequence

$$F^j = 4^j f^j / \|f^j\|_p,$$

which satisfies the hypothesis of the theorem since

$$L(F^j) = (4^j / \|f^j\|_p) L(f^j),$$

which is certainly bounded. Clearly $\|F^j\|_p = 4^j$ and our next step is to derive a contradiction from this fact by constructing an L for which the sequence $L(F^j)$ is not bounded.

Set $T_j(x) = |F^j(x)|^{p-2}\overline{F^j}(x)/\|F^j\|_p^{p-1}$ and define complex numbers σ_n of modulus 1 as follows: pick $\sigma_1 = 1$ and choose σ_n recursively by requiring $\sigma_n \int T_n F^n$ to have the same argument as

$$\sum_{j=1}^{n-1} 3^{-j} \sigma_j \int T_j F^n.$$

Thus,

$$\left| \sum_{j=1}^{n} 3^{-j} \sigma_j \int T_j F^n \right| \geq 3^{-n} \int T_n F^n = 3^{-n} \|F^n\|_p = (4/3)^n.$$

Now define the linear functional L by setting

$$L(h) = \sum_{j=1}^{\infty} 3^{-j} \sigma_j \int T_j h,$$

which is obviously continuous by Hölder's inequality and the fact that $\|T_j\|_{p'} = 1$.

We can bound $|L(F^k)|$ from below as follows.

$$|L(F^k)| \geq \left| \sum_{j=1}^{k} 3^{-j} \sigma_j \int T_j F^k \right| - \left(\sum_{j=k+1}^{\infty} 3^{-j} \right) 4^k$$

$$\geq 3^{-k} 4^k - 3^{-k} 4^k \frac{1/3}{1 - (1/3)} = \frac{1}{2} \left(\frac{4}{3} \right)^k$$

which tends to ∞ as $k \to \infty$. This contradicts the boundedness of $L(F^k)$. ■

● The next theorem, [Mazur], shows how to build strongly convergent sequences out of weakly convergent ones. It can be very useful for proving existence of minimizers for variational problems. In fact, we shall employ it in the capacitor problem in Chapter 11. The theorem holds in greater generality than the version we give here, e.g., it also holds for $L^1(\Omega)$ and $L^\infty(\Omega)$. In fact it holds for any normed space (see [Rudin 1991], Theorem 3.13). We prove it for $1 < p < \infty$ by using Lemma 2.8 (projection on convex sets). For full generality it is necessary to use the Hahn–Banach theorem, which involves the axiom of choice and which the reader can find in many texts. The proof here is somewhat more constructive and intuitive.

2.13 THEOREM (Strongly convergent convex combinations)

Let $1 < p < \infty$ and let $f^1, f^2, \dots$ be a sequence in $L^p(\Omega)$ that converges weakly to $F \in L^p(\Omega)$. Then we can form a sequence $F^1, F^2, \dots$ in $L^p(\Omega)$ that converges strongly *to F, and such that each F^j is a convex combination of the functions $f^1, \dots, f^j$. I.e., for each j there are nonnegative numbers $c_1^j, \dots, c_j^j$ such that $\sum_{k=1}^j c_k^j = 1$ and such that the functions*

$$F^j := \sum_{k=1}^{j} c_k^j f^k$$

converge strongly to F.

PROOF. First, consider the set $\widetilde{K} \subset L^p(\Omega)$ which consists of all the f^j's together with all finite convex combinations of them, i.e., all functions of the form $\sum_{k=1}^m d_k f^k$ with m arbitrary and with $\sum_{k=1}^m d_k = 1$ where $d_k \geq 0$. This set $\widetilde{K}$ is clearly convex, i.e. $f, g \in \widetilde{K} \Rightarrow \lambda f + (1-\lambda) g \in \widetilde{K}$ for all $0 \leq \lambda \leq 1$.

Next, let K denote the union of $\widetilde{K}$ and all its limit points, i.e. we add to $\widetilde{K}$ all functions in $L^p(\Omega)$ that are limits of Cauchy sequences of elements of $\widetilde{K}$. We claim that (a) K is convex and (b) K is closed. To prove (a) we note that if $f^j \to f$ and $g^j \to g$ (with $f^j, g^j \in \widetilde{K}$) then $\lambda f^j + (1-\lambda) g^j \in \widetilde{K}$ and converges to $\lambda f + (1-\lambda) g$. To prove (b), the reader can use the triangle inequality to prove that 'Cauchy sequences of Cauchy sequences are Cauchy sequences'. (Our construction here imitates the construction of the reals from the rationals.)

Our theorem amounts to the assertion that the weak limit F is in K. Suppose otherwise. By Lemma 2.8 (projection on convex sets) there is a function $h \in K$ such that $D = \text{dist}(F, K) = \|F - h\|_p > 0$. In 2.8(2) we considered the function

$$\ell(x) = [\overline{F}(x) - \overline{h}(x)]|F(x) - h(x)|^{p-2}$$

which is in $L^{p'}(\Omega)$ and showed that the continuous linear function $L(g) := \int \ell g$ satisfies

$$\operatorname{Re} L(g) - \operatorname{Re} L(h) \le 0 \tag{1}$$

for all $g \in K$. However, $L(F - h) = \|F - h\|_p^p$, and hence

$$\operatorname{Re} L(F) - \operatorname{Re} L(h) > 0 \tag{2}$$

because $F - h$ is not the zero function. (2) contradicts (1) because $L(f^j) \to L(F)$ by assumption, and the f^j's are in K. ■

● At last we come to the identification of $L^p(\Omega)^*$, the dual of $L^p(\Omega)$, for $1 \le p < \infty$. This is **F. Riesz's representation theorem**. The dual of $L^\infty(\Omega)$ is not given because it is a huge, less useful space that requires the axiom of choice for its construction.

2.14 THEOREM (The dual of $L^p(\Omega)$)

When $1 \le p < \infty$ the dual of $L^p(\Omega)$ is $L^q(\Omega)$, with $1/p + 1/q = 1$, in the sense that every $L \in L^p(\Omega)^*$ has the form

$$L(g) = \int_\Omega v(x) g(x) \mu(\mathrm{d}x) \tag{1}$$

for some unique $v \in L^q(\Omega)$. (In case $p = 1$ we make the additional technical assumption that (Ω, μ) is sigma-finite.) In all cases, even $p = \infty$, L given by (1) *is in $L^p(\Omega)^*$ and its norm (defined in* 2.9(5)*) is*

$$\|L\| = \|v\|_q. \tag{2}$$

PROOF. $\boxed{1 < p < \infty:}$ With $L \in L^p(\Omega)^*$ given, define the set $K = \{g \in L^p(\Omega) : L(g) = 0\} \subset L^p(\Omega)$. Clearly K is convex and K is closed (here is where the continuity of L enters). Assume $L \neq 0$, whence there is $f \in L^p(\Omega)$ such that $L(f) \neq 0$, i.e., $f \notin K$. By Lemma 2.8 (projection on convex sets) there is an $h \in K$ such that

$$\operatorname{Re} \int uk \leq 0 \tag{3}$$

for all $k \in K$. Here $u(x) = |f(x) - h(x)|^{p-2}[\overline{f}(x) - \overline{h}(x)]$, which is evidently in $L^q(\Omega)$. However, K is a linear space and hence $-k \in K$ and $ik \in K$ whenever $k \in K$. The first fact tells us that $\operatorname{Re} \int uk = 0$ and the second fact implies $\int uk = 0$ for all $k \in K$.

Now let g be an arbitrary element of $L^p(\Omega)$ and write $g = g_1 + g_2$ with

$$g_1 = \frac{L(g)}{L(f-h)}(f-h) \quad \text{and} \quad g_2 = g - g_1.$$

(Note that $L(f-h) = L(f) \neq 0$.) One easily checks that $L(g_2) = 0$, i.e., $g_2 \in K$, whence

$$\int ug = \int ug_1 + \int ug_2 = \int ug_1 = L(g)A,$$

where $A = \int u(f-h)/L(f-h) \neq 0$, since $\int u(f-h) = \int |f-h|^p$. Thus, the v in (1) equals u/A. The uniqueness of v follows from the fact that if $\int (v-w)g = 0$ for all $g \in L^p(\Omega)$, and with $w \in L^q(\Omega)$, then we could obtain a contradiction by choosing $g = (\overline{v-w})|v-w|^{q-2} \in L^p(\Omega)$. The easy proof of (2) is left to the reader.

$\boxed{p = 1:}$ Let us assume for the moment that Ω has finite measure. In this case, Hölder's inequality implies that a continuous linear functional L on $L^1(\Omega)$ has a restriction to $L^p(\Omega)$ which is again continuous since

$$|L(f)| \leq C\|f\|_1 \leq C\mu(\Omega)^{1/q}\|f\|_p \tag{4}$$

for all $p \geq 1$. By the previous proof for $p > 1$, we have the existence of a unique $v_p \in L^q(\Omega)$ such that $L(f) = \int v_p(x)f(x)\mu(\mathrm{d}x)$ for all $f \in L^p(\Omega)$. Moreover, since $L^r(\Omega) \subset L^p(\Omega)$ for $r \geq p$ (by Hölder's inequality) the uniqueness of v_p for each p implies that v_p is, in fact, independent of p, i.e., this function (which we now call v) is in every $L^r(\Omega)$-space for $1 < r < \infty$.

If we now pick some dual pair q and p with $p > 1$ and choose $f = |v|^{q-2}\overline{v}$ in (4) we obtain

$$\int |v|^q = L(f) \leq C(\mu(\Omega))^{1/q}\left(\int |v|^{(q-1)p}\right)^{1/p} = C(\mu(\Omega))^{1/q}\|v\|_q^{q-1},$$

and hence $\|v\|_q \leq C(\mu(\Omega))^{1/q}$ for all $q < \infty$. We claim that $v \in L^\infty(\Omega)$; in fact $\|v\|_\infty \leq C$. Suppose that $\mu(\{x \in \Omega : |v(x)| > C + \varepsilon\}) = M > 0$. Then $\|v\|_q \geq (C+\varepsilon)M^{1/q}$, which exceeds $C\mu(\Omega)^{1/q}$ if q is big enough.

Thus $v \in L^\infty(\Omega)$ and $L(f) = \int v(x)f(x)\,\mathrm{d}\mu$ for all $f \in L^p(\Omega)$ for any $p > 1$. If $f \in L^1(\Omega)$ is given, then $\int |v(x)||f(x)|\,\mathrm{d}\mu < \infty$. Replacing $f(x)$ by $f^k(x) = f(x)$ whenever $|f(x)| \le k$ and by zero otherwise, we note that $|f^k(x)| \le |f(x)|$ and $f^k(x) \to f(x)$ pointwise as $k \to \infty$; hence, by dominated convergence, $f^k \to f$ in $L^1(\Omega)$ and $vf^k \to vf$ in $L^1(\Omega)$. Thus

$$L(f) = \lim_{k\to\infty} L(f^k) = \lim_{k\to\infty} \int vf^k\,\mathrm{d}\mu = \int vf\,\mathrm{d}\mu.$$

The previous conclusion can be extended to the case that $\mu(\Omega) = \infty$ but Ω is sigma-finite. Then

$$\Omega = \bigcup_{j=1}^{\infty} \Omega_j$$

with $\mu(\Omega_j)$ finite and with $\Omega_j \cap \Omega_k$ empty whenever $j \ne k$. Any $L^1(\Omega)$ function f can be written as

$$f(x) = \sum_{j=1}^{\infty} f_j(x)$$

where $f_j = \chi_j f$ and χ_j is the characteristic function of Ω_j. $f_j \mapsto L(f_j)$ is then an element of $L^1(\Omega_j)^*$, and hence there is a function $v_j \in L^\infty(\Omega_j)$ such that $L(f_j) = \int_{\Omega_j} v_j f_j = \int_{\Omega_j} v_j f$. The important point is that each v_j is bounded in $L^\infty(\Omega_j)$ by the *same* $C = \|L\|$. Moreover, the function v, defined on all of Ω by $v(x) = v_j(x)$ for $x \in \Omega_j$, is clearly measurable and bounded by C. Thus, we have $L(f) = \int_\Omega vf$ by the countable additivity of the measure μ. Uniqueness is left to the reader. ■

● Our next goal is the Banach–Alaoglu Theorem, 2.18, and, although it can be presented in a much more general setting, we restrict ourselves to the particular case in which Ω is a subset of $\mathbb{R}^n$ and $\mu(\mathrm{d}x)$ is Lebesgue measure. To reach it we need the separability of $L^p(\Omega)$ for $1 < p < \infty$ and to achieve that we need the density of continuous functions in $L^p(\Omega)$. The next theorem establishes this fact, and it is one of the most fundamental; its importance cannot be overstressed. It permits us to approximate $L^p(\Omega)$, functions by C_c^∞-functions (Lemma 2.19). Why then, the reader might ask, did we introduce the L^p-spaces? Why not restrict ourselves to the C^∞-functions from the outset? The answer is that the set of continuous functions is not complete in $L^p(\Omega)$, i.e., the analogue of Theorem 2.7 does not hold for them because limits of continuous functions are not necessarily continuous. As preparation we need 2.15–2.17.

2.15 CONVOLUTION

When f and g are two (complex-valued) functions on $\mathbb{R}^n$ we define their *convolution* to be the function $f * g$ given by

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y. \tag{1}$$

Note that $f * g = g * f$ by changing variables. One has to be careful to make sure that (1) makes sense. One way is to require $f \in L^p(\mathbb{R}^n)$ and $g \in L^{p'}(\mathbb{R}^n)$, in which case the integral in (1) is well defined for *all* x by Hölder's inequality. More is true, as Lemma 2.20 and Theorem 4.2 (Young's inequality) show. In case f and g are in $L^1(\mathbb{R}^n)$, (1) makes sense for almost every $x \in \mathbb{R}^n$ and defines a measurable function that is in $L^1(\mathbb{R}^n)$ (see Exercise 7). Indeed, Theorem 4.2 shows that when $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ with $1/p+1/q \geq 1$, then (1) is finite a.e. and defines a measurable function that is in $L^r(\mathbb{R}^n)$ with $1+1/r = 1/p+1/q$. In the following theorem we prove this for $q = 1$.

2.16 THEOREM (Approximation by C^∞-functions)

Let j be in $L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} j = 1$. For $\varepsilon > 0$, define $j_\varepsilon(x) := \varepsilon^{-n} j(x/\varepsilon)$, so that $\int_{\mathbb{R}^n} j_\varepsilon = 1$ and $\|j_\varepsilon\|_1 = \|j\|_1$. Let $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p < \infty$ and define the convolution

$$f_\varepsilon := j_\varepsilon * f.$$

Then

$$f_\varepsilon \in L^p(\mathbb{R}^n) \quad \textit{and} \quad \|f_\varepsilon\|_p \leq \|j\|_1 \|f\|_p. \tag{1}$$

$$f_\varepsilon \to f \quad \textit{strongly in} \quad L^p(\mathbb{R}^n) \ \textit{as}\ \varepsilon \to 0. \tag{2}$$

If $j \in C_c^\infty(\mathbb{R}^n)$, then $f_\varepsilon \in C^\infty(\mathbb{R}^n)$ and (see Remark (3) below)

$$D^\alpha f_\varepsilon = (D^\alpha j_\varepsilon) * f. \tag{3}$$

REMARKS. (1) The above theorem is stated for $\mathbb{R}^n$ but it applies equally well to any measurable set $\Omega \subset \mathbb{R}^n$. Given $f \in L^p(\Omega)$ we can define $\widetilde{f} \in L^p(\mathbb{R}^n)$ by $\widetilde{f}(x) = f(x)$ for $x \in \Omega$ and $\widetilde{f}(x) = 0$ for $x \notin \Omega$. Then define

$$f_\varepsilon(x) = (j_\varepsilon * \widetilde{f})(x) \quad \text{for } x \in \Omega.$$

Equation (1) holds in $L^p(\Omega)$ since

$$\|f_\varepsilon\|_{L^p(\Omega)} \leq \|f_\varepsilon\|_{L^p(\mathbb{R}^n)} \leq \|j\|_1 \|\widetilde{f}\|_{L^p(\mathbb{R}^n)} = \|j\|_1 \|f\|_{L^p(\Omega)}.$$

Likewise, (2) is correct in $L^p(\Omega)$. If Ω is open (so that $C^\infty(\Omega)$ can be defined), then the third statement obviously holds as well with $C^\infty(\mathbb{R}^n)$ replaced by $C^\infty(\Omega)$ and f replaced by $\widetilde{f}$.

(2) We shall see in Lemma 2.19 that Theorem 2.16 can be extended in another way: The $C^\infty(\mathbb{R}^n)$ approximants, $j_\varepsilon * f$, can be modified so that they are in $C_c^\infty(\mathbb{R}^n)$ without spoiling conclusions (1) and (2). The proof of Lemma 2.19 is an easy exercise, but the lemma is stated separately because of its importance.

(3) In Chapter 6 we shall define the distributional derivative of an L^p function, f, denoted by $D^\alpha f$. It is then true that $(D^\alpha j_\varepsilon) * f = j_\varepsilon * D^\alpha f$.

(4) In Theorem 1.19 (approximation by C^∞ functions) we proved that any $f \in L^1(\mathbb{R}^n)$ can be approximated (in the $L^1(\mathbb{R}^n)$ norm) by $C^\infty(\mathbb{R}^n)$ functions. One of our purposes here is to be more explicit by showing that $C^\infty(\mathbb{R}^n)$ can be generated by convolution. This is not our only concern, however; statement (2) will also be important later. Theorem 1.18 (approximation by really simple functions) will play a key role in our proof.

PROOF. Statement (1) is Young's inequality, which will be proved in Sect 4.2. Only the "simple version" proved in part (A) of the proof, is needed, i.e., 4.2(4), but with $C_{p',q,r;n}$ replaced by 1. This version is only a simple exercise using Hölder's inequality. We shall use it freely in our proof here and ask the readers's indulgence for this forward leap to Chapter 4.

To prove (2) we have to show that for every $\delta > 0$ we can find an $\varepsilon > 0$ such that $\|f_\varepsilon - f\|_p < 10\delta$.

Step 1. We claim that we may assume that j and f have compact support and that $|f|$ is bounded, i.e., $f \in L^\infty(\mathbb{R}^n)$. If j does not have compact support we can (by dominated convergence) find $0 < R < \infty$ and $C > 1$ such that $j^R(x) := C\chi_{\{|x|<R\}}(x)j(x)$ satisfies $\int_{\mathbb{R}^n} j^R = 1$ and $\|f\|_p\|j - j^R\|_1 < \delta$. Define $j_\varepsilon^R = \varepsilon^{-n} j^R(x/\varepsilon)$ (which has support in $\{x : |x| < R\varepsilon\}$), and note that the number $\|j_\varepsilon - j_\varepsilon^R\|_1$ is independent of ε. By Young's inequality, $\|j_\varepsilon * f - j_\varepsilon^R * f\|_p = \|(j_\varepsilon - j_\varepsilon^R) * f\|_p < \delta$. By the triangle inequality, if we can prove that $\|j_\varepsilon^R * f - f\|_p < \delta$ for small enough ε we will have that $\|j_\varepsilon * f - f\|_p < 2\delta$. Henceforth, we shall omit the R and just assume that j has support in a ball of radius R.

In a similar fashion, to within an error 2δ we can replace $f(x)$ by $\chi_{\{|x|<R'\}}(x)f(x)$ for some sufficiently large R'. The compact support of f implies that $f \in L^1(\mathbb{R}^n)$; in fact, $\|f\|_1 \le (|\mathbb{S}^{n-1}|/n)(R')^{n/p'}\|f\|_p$.

Using Young's inequality and dominated convergence once again we can also replace $f(x)$ by the cut off function $\chi_{\{|f|<h\}}(x)f(x)$ for some sufficiently

large h at the cost of an additional error δ. The fact that now $\|f\|_\infty \le h$ implies that $\|j_\varepsilon * f\|_\infty \le h$ and that

$$\|j_\varepsilon * f - f\|_p \le (2h)^{1/p'} \|j_\varepsilon * f - f\|_1.$$

Our conclusion in this first step is the following: To prove (2) it suffices to assume that j has support in a ball of radius R and to assume that $p = 1$. We shall now prove (2) under these conditions.

Step 2. By Theorem 1.18 there is a really simple function F (using the algebra of half open rectangles in 1.17(1)) such that $\|F - f\|_1 < \delta$, and hence (by Young's inequality) $\|j_\varepsilon * F - j_\varepsilon * f\|_1 < \delta$. By the triangle inequality, it suffices to prove that $\|j_\varepsilon * F - F\| < \delta$ for sufficiently small ε, but since F is just a finite linear combination of characteristic functions of rectangles (say, N of them) it suffices to show that for every rectangle H

$$\lim_{\varepsilon \to 0} \|j_\varepsilon * \chi_H - \chi_H\|_1 = 0, \tag{4}$$

where χ_H is the characteristic function of H. (As far as (4) is concerned it does not matter whether H is closed or open.)

Recall that j_ε has support in a ball of radius $r = R\varepsilon$ and this r can be made as small as we please. We choose r so small that the sets $A_- = \{x \in H : \text{distance}(x, H^c) < r\}$ and $A_+ = \{x \notin H : \text{distance}(x, H) < r\}$ satisfy $\mathcal{L}^n(A_- \cup A_+) < \delta/\|j\|_1$. Clearly, if $x \notin A_- \cup A_+$, then $j_\varepsilon * \chi_H(x) = \chi_H(x)$ since $\int_{\mathbb{R}^n} j = 1$. If $x \in A_- \cup A_+$, then

$$|j_\varepsilon * \chi_H(x) - \chi_H(x)| = \left| \int_{\mathbb{R}^n} j(y)[H(x-y) - H(x)]\mathrm{d}y \right| \le \int_{\mathbb{R}^n} |j|.$$

Since $\mathcal{L}^n(A_- \cup A_+) < \delta/\|j\|_1$, this proves (2).

Step 3. To prove (3) we shall prove that

$$\partial f_\varepsilon / \partial x_i = (\partial j_\varepsilon / \partial x_i) * f, \tag{5}$$

and that this function is continuous. This will imply that $f_\varepsilon \in C^1(\mathbb{R}^n)$ and, by induction (since $\partial j_\varepsilon/\partial x_i \in C^\infty(\mathbb{R}^n)$), that $f_\varepsilon \in C^\infty(\mathbb{R}^n)$. The continuity is an elementary consequence of the dominated convergence theorem. Since the support of j_ε is compact, the difference quotient

$$\Delta_{\varepsilon,\delta}(x) := [j_\varepsilon(\dots, x_i + \delta, \dots) - j_\varepsilon(\dots, x_i, \dots)]/\delta$$

is uniformly bounded in δ and of compact support and it is obviously bounded by some fixed $L^{p'}$-function. The desired conclusion follows again by dominated convergence. ■

2.17 LEMMA (Separability of $L^p(\mathbb{R}^n)$)

There exists a fixed, countable set of functions $\mathcal{F} = \{\phi_1, \phi_2, \dots\}$ (which will be constructed explicitly) with the following property: For each $1 \le p < \infty$ and for each measurable set $\Omega \subset \mathbb{R}^n$, for each function $f \in L^p(\Omega)$ and for each $\varepsilon > 0$ we have $\|f - \phi_j\|_p < \varepsilon$ for some function ϕ_j in $\mathcal{F}$.

REMARK. The separability of $L^1(\Omega)$ is an immediate consequence of Theorem 1.18, using the algebra generated by the half open rectangles 1.17(1). This can be easily extended to $L^p(\Omega)$ for general p. The proof below, however, yields a useful and fairly explicit construction of the family $\mathcal{F}$.

PROOF. It suffices to prove this for $\Omega = \mathbb{R}^n$ since we always can extend $f \in L^p(\Omega)$ to a function in $L^p(\mathbb{R}^n)$ by setting $f(x) = 0$ for $x \notin \Omega$.

To define $\mathcal{F}$ we first define a countable family, Γ, of sets in $\mathbb{R}^n$ as the collection of cubes $\Gamma_{j,m}$, for $j = 1, 2, 3, \dots$ and for $m \in \mathbb{Z}^n$, given by

$$\Gamma_{j,m} = \{x \in \mathbb{R}^n : 2^{-j} m_i < x_i \le 2^{-j}(m_i + 1),\ i = 1, \dots, n\}.$$

For each j, the $\Gamma_{j,m}$'s obviously cover the whole of $\mathbb{R}^n$ as m ranges over $\mathbb{Z}^n$, the points in $\mathbb{R}^n$ with integer coordinates. The family Γ is a countable family (here we use the fact that a countable family of countable families is countable).

Next, we define the family of functions $\mathcal{F}_j$ to consist of all functions f on $\mathbb{R}^n$ with the property that $f(x) = c_{j,m} =$ constant for $x \in \Gamma_{j,m}$ and, moreover, the numbers $c_{j,m}$ are restricted to be rational complex numbers. Again this family $\mathcal{F}_j$ is countable. $\mathcal{F}$ is defined to be $\bigcup_{j=1}^{\infty} \mathcal{F}_j$, which is again countable.

Given $f \in L^p(\mathbb{R}^n)$, we first use Theorem 2.16 to replace f by a continuous function $\widetilde{f} \in L^p(\mathbb{R}^n)$ such that $\int |f - \widetilde{f}|^p < \varepsilon/3$. Thus, it suffices to find $f_j \in \mathcal{F}$ such that $\int |\widetilde{f} - f_j|^p < 2\varepsilon/3$. We can also assume (as in the proof of 2.16) that $\widetilde{f}(x) = 0$ for x outside some large cube γ of the form $\{x : -2^J \le x_i < 2^J\}$ for some integer J.

For each integer j we define

$$\widetilde{f}_j(x) = 2^{-nj} \int_{\Gamma_{j,m}} \widetilde{f}(y)\, \mathrm{d}y \quad \text{for } x \in \Gamma_{j,m},$$

i.e., $\widetilde{f}_j$ is the average of $\widetilde{f}$ over $\Gamma_{j,m}$. Since $\widetilde{f}$ is continuous, it is uniformly continuous on γ. This means that for each $\varepsilon' > 0$ there is a $\delta > 0$ such that $|\widetilde{f}(y) - \widetilde{f}(x)| < \varepsilon'$ whenever $|x - y| < \delta$. Therefore, if j is large enough so that $\delta \geq \sqrt{n}2^{-j}$, we have

$$\int_{\mathbb{R}^n} |\widetilde{f}(x) - \widetilde{f}_j(x)|^p \, \mathrm{d}x \leq \text{volume}(\gamma)(2\varepsilon')^p.$$

We can choose ε' to satisfy $(2\varepsilon')^p$ volume$(\gamma) < \varepsilon/3$. Thus, $\int |f - \widetilde{f}_j|^p < \varepsilon/3$.

The final step is to replace $\widetilde{f}_j$ by a function $\widehat{f}_j$ that assumes only rational complex values in such a way that $\int |\widetilde{f}_j - \widehat{f}_j|^p < \varepsilon/3$. This is easy to do since only finitely many cubes (and hence only finitely many values of $\widetilde{f}_j$) are involved. Since $\widehat{f}_j \in \mathcal{F}$, our goal has been accomplished. ■

● The next theorem is the **Banach–Alaoglu theorem**, but for the special case of L^p-spaces. As such, it predates Banach–Alaoglu (although we shall continue to use that appellation). For the case at hand, i.e., L^p-spaces, the axiom of choice in the realm of the uncountable is *not* needed in the proof.

2.18 THEOREM (Bounded sequences have weak limits)

Let $\Omega \in \mathbb{R}^n$ be a measurable set and consider $L^p(\Omega)$ with $1 < p < \infty$. Let $f^1, f^2, \dots$ be a sequence of functions, bounded in $L^p(\Omega)$. Then there exist a subsequence $f^{n_1}, f^{n_2}, \dots$ (with $n_1 < n_2 < \cdots$) and an $f \in L^p(\Omega)$ such that $f^{n_i} \rightharpoonup f$ weakly in $L^p(\Omega)$ as $i \to \infty$, i.e., for every bounded linear functional $L \in L^p(\Omega)^*$

$$L(f^{n_i}) \to L(f) \quad \text{as} \quad i \to \infty.$$

PROOF. We know from Riesz's representation theorem, Theorem 2.14, that the dual of $L^p(\Omega)$ is $L^q(\Omega)$ with $1/p + 1/q = 1$. Therefore, our first task is to find a subsequence f^{n_j} such that $\int f^{n_j}(x)g(x)\,\mathrm{d}\mu$ is a convergent sequence of numbers for every $g \in L^q(\Omega)$. In view of Lemma 2.17 (separability of $L^p(\mathbb{R}^n)$), it suffices to show this convergence only for the special countable sequence of functions ϕ^j given there.

Cantor's diagonal argument will be used. First, consider the sequence of numbers $C_1^j = \int f^j \phi_1$, which is bounded (by Hölder's inequality and the boundedness of $\|f^j\|_p$). There is then a subsequence (which we denote by f_1^j) such that C_1^j converges to some number C_1 as $j \to \infty$. Second, starting with this new sequence $f_1^1, f_1^2, \dots$, a parallel argument shows that we can pass to a further subsequence such that $C_2^j = \int f^j \phi_2$ also converges to some number C_2. This second subsequence is denoted by $f_2^1, f_2^2, f_2^3, \dots$. Proceeding inductively we generate a countable family of subsequences so that for the k^{th} subsequence (and all further subsequences) $\int f_k^j \phi_k$ converges as $j \to \infty$. Moreover, f_ℓ^j is somewhere in the sequence $f_k^1, f_k^2, \ldots$ if $k \le \ell$.

Cantor told us how to construct one convergent subsequence from all these. The k^{th} function in this new sequence f^{n_k} (which will henceforth be called F^k) is defined to be the k^{th} function in the k^{th} sequence, i.e., $F^k := f_k^k$. It is a simple exercise to show that $\int F^k \phi_\ell \to C_\ell$ as $j \to \infty$.

Our second and final task is to use the knowledge that $\int F^j g$ converges to some number (call it $L(g)$) as $j \to \infty$ for all $g \in L^q(\mathbb{R}^n)$ in order to show the existence of an $f \in L^p$ to which F^j converges weakly. To do so we note that $L(g)$ is clearly a linear functional on $L^q(\mathbb{R}^n)$ and it is also bounded (and hence continuous) since $\|F^j\|_p$ is bounded. But Theorem 2.14 tells us that the dual of $L^q(\mathbb{R}^n)$ is precisely $L^p(\mathbb{R}^n)$, and hence there is some $f \in L^p(\mathbb{R}^n)$ such that $\int F^j g \to L(g) = \int fg$. ■

REMARK. What was really used here was the fact that the 'double dual' (or the 'dual of the dual') of $L^p(\mathbb{R}^n)$ is $L^p(\mathbb{R}^n)$. For other spaces, such as $L^1(\mathbb{R}^n)$ or $L^\infty(\mathbb{R}^n)$, the double dual is larger than the starting space, and then the analogue of Theorem 2.18 fails. Here is a counterexample in $L^1(\mathbb{R}^1)$. Let $f^j(x) = j$ for $0 \le x \le 1/j$ and zero otherwise. This sequence is certainly bounded: $\int |f^j| = 1$. If some subsequence had a weak limit, f, then f would have to be zero (because f would have to be zero on all intervals of the form $(-\infty, 0)$ or $(1/n, \infty)$ for any n. But $\int f^j \cdot 1 = 1 \not\to 0$, which is a contradiction since the function $f(x) \equiv 1$ is in the dual space $L^\infty(\mathbb{R}^1)$.

2.19 LEMMA (Approximation by C_c^∞-functions)

Let $\Omega \subset \mathbb{R}^n$ be an open set and let $K \subset \Omega$ be compact. Then there is a function $J_K \in C_c^\infty(\Omega)$ such that $0 \le J_K(x) \le 1$ for all $x \in \Omega$ and $J_K(x) = 1$ for $x \in K$.

As a consequence, there is a sequence of functions $g_1, g_2, \ldots$ in $C_c^\infty(\Omega)$ that take values in $[0,1]$ and such that $\lim_{j\to\infty} g_j(x) = 1$ for every $x \in \Omega$.

As a second consequence, given any sequence of functions $f_1, f_2, \ldots$ in $C^\infty(\Omega)$ that converges strongly to some f in $L^p(\Omega)$ with $1 \le p < \infty$, the sequence given by $h_i(x) = g_i(x) f_i(x)$ is in $C_c^\infty(\Omega)$ and also converges to f in the same strong sense. If, on the other hand, $f_i \rightharpoonup f$ weakly in $L^p(\mathbb{R}^n)$ for some $1 < p < \infty$, then $h_i \rightharpoonup f$ weakly in $L^p(\mathbb{R}^n)$.

PROOF. The first part of Lemma 2.19 is Urysohn's Lemma (Exercise 1.15) but we shall give a short proof using the *Lebesgue* integral instead of the

Riemann integral. Since K is compact, there is a $d > 0$ such that $\{x : |x-y| \le 2d$ for some $y \in K\} \subset \Omega$. Define $K_+ = \{x : |x-y| \le d$ for some $y \in K\} \supset K$ and note that $K_+ \subset \Omega$ is also compact. Fix some $j \in C_c^\infty(\mathbb{R}^n)$ with support in $\{x : |x| \le 1\}$ and such that $0 \le j(x) \le 1$ for all x and $\int j = 1$ (see 1.1(2) for an example). Then, with $\varepsilon = d$, we set $J_K = j_\varepsilon * \chi$, where χ is the characteristic function of K_+. It is evident that J_K has the correct properties.

It is an easy exercise to show that there is an increasing sequence of compact sets $K_1 \subset K_2 \subset \cdots \subset \Omega$ such that each $x \in \Omega$ is in $K_{m(x)}$ for some integer $m(x)$. Define $g_i := J_{K_i}$.

The strong convergence of h_i to f is a consequence of dominated convergence. The weak convergence is also a consequence of dominated convergence provided we recall that the dual of $L^p(\Omega)$ is $L^{p'}(\Omega)$, with $1 < p' < \infty$, and that the functions of compact support are dense in $L^{p'}(\Omega)$. ■

2.20 LEMMA (Convolutions of functions in dual $L^p(\mathbb{R}^n)$-spaces are continuous)

Let f be a function in $L^p(\mathbb{R}^n)$ and let g be in $L^{p'}(\mathbb{R}^n)$ with p and $p' > 1$ and $1/p + 1/p' = 1$. Then the convolution $f * g$ is a continuous function on $\mathbb{R}^n$ that tends to zero at infinity in the strong sense that for any $\varepsilon > 0$ there is $\mathcal{R}_\varepsilon$ such that

$$\sup_{|x| > \mathcal{R}_\varepsilon} |(f*g)(x)| < \varepsilon.$$

PROOF. Note that $(f*g)(x)$ is finite and defined by $\int f(x-y)g(y)\,\mathrm{d}y$ for *every* x. This follows from Hölder's inequality since $f \in L^p(\mathbb{R}^n)$ and $g \in L^{p'}(\mathbb{R}^n)$. For any $\delta > 0$ we can find, by Lemma 2.19 (approximation by $C_c^\infty(\Omega)$-functions), f_δ and g_δ, both in $C_c^\infty(\mathbb{R}^n)$, such that $\|f_\delta - f\|_p \le \delta$ and $\|g_\delta - g\|_{p'} \le \delta$. If we write

$$f * g - f_\delta * g_\delta = (f - f_\delta) * g + f_\delta * (g - g_\delta),$$

we see, by the triangle and Hölder's inequalities, that

$$\|f * g - f_\delta * g_\delta\|_\infty \le \|f - f_\delta\|_p \|g\|_{p'} + \|f_\delta\|_p \|g - g_\delta\|_{p'},$$

which is bounded by $(\|g\|_{p'} + \|f\|_p)\delta$. Since $f_\delta * g_\delta$ is in $C_c^\infty(\mathbb{R}^n)$, $f * g$ is uniformly approximated by smooth functions. Hence $f * g$ is continuous and the last statement is a trivial consequence of the fact that $f_\delta * g_\delta$ has compact support. ■

2.21 HILBERT-SPACES

The space $L^2(\Omega)$ has the special property, not shared by the other L^p-spaces, that its norm is given by an inner product—a concept familiar from elementary linear algebra. The inner product of two $L^2(\Omega)$ functions is

$$(f, g) := \int_\Omega \overline{f}(x) g(x) \mu(\mathrm{d}x)$$

in terms of which the norm is given by $\|f\|^2 = \sqrt{(f, f)}$. Note that the complex conjugate is on the left; often it is on the right, especially in mathematical writing. Note also that the function $\overline{f}g$ is integrable, by Schwarz's inequality.

Hilbert-spaces can be defined abstractly in terms of the inner product, without mentioning functions, similar to the way a vector space can be defined without any specific representation of the vectors. In this section we shall outline the beginning of that theory.

Generally speaking, an **inner product space** V is a vector space that carries an **inner product** $(\cdot, \cdot) : V \times V \to \mathbb{C}$ having the properties

(i) $(x, y + z) = (x, y) + (x, z)$ for all $x, y, z \in V$;
(ii) $(x, \alpha y) = \alpha(x, y)$ for all $x, y \in V, \alpha \in \mathbb{C}$;
(iii) $(y, x) = \overline{(x, y)}$;
(iv) $(x, x) \geq 0$ for all x, and $(x, x) = 0$ only if $x = 0$.

Clearly, $\int \overline{f} g \,\mathrm{d}\mu$ satisfies all these conditions.

The Schwarz inequality $|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}$ can now be deduced from (i)–(iv) alone. If one of the vectors, say y, is not the zero vector, then there is equality if and only if $x = \lambda y$ for some $\lambda \in \mathbb{C}$. As an exercise the reader is asked to prove this. If we set $\|x\| = \sqrt{(x, x)}$, then, by the Schwarz inequality,

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

and hence the triangle inequality $\|x + y\| \leq \|x\| + \|y\|$ holds. With the help of (ii) and (iv) the function $x \mapsto \|x\|$ is seen to be a norm.

We say that $x, y \in V$ are **orthogonal** if $(x, y) = 0$. Keeping with the tradition that every deep theorem becomes trivial with the right definition, we can state **Pythagoras' theorem** in the following way: When x and y are orthogonal, $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

An important property of $L^2(\Omega)$ is its completeness. A **Hilbert-space** $\mathcal{H}$ is *by definition* a complete inner product space, i.e., for every Cauchy sequence $x^j \in \mathcal{H}$ (meaning that $\|x^j - x^k\| \to 0$ as $j, k \to \infty$) there is some $x \in \mathcal{H}$ such that $\|x - x^j\| \to 0$ as $j \to \infty$.

With these preparations, we invite the reader to prove, as an exercise, the analogue of Lemma 2.8 (projection on convex sets) for Hilbert-spaces: Let $\mathcal{C}$ be a closed convex set in $\mathcal{H}$. Then there exists an element y of smallest norm in $\mathcal{C}$, i.e., such that $\|y\| = \inf\{\|x\| : x \in \mathcal{C}\}$.

The uniform convexity, which is needed for the projection lemma, is provided by the parallelogram identity

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

As in Theorem 2.14, the projection lemma implies that the dual of $\mathcal{H}$, i.e., the continuous linear functionals on $\mathcal{H}$, is $\mathcal{H}$ itself.

A special case of a convex set is a **subspace** of a Hilbert-space $\mathcal{H}$, i.e., a set $M \subset \mathcal{H}$ that is closed under finite linear combinations. Let $M^\perp$ be the **orthogonal complement** of M, i.e.,

$$M^\perp := \{x \in \mathcal{H} : (x,y) = 0, y \in M\}.$$

It is easy to see that $M^\perp$ is a **closed subspace**, i.e., if $x^j \in M^\perp$ and $x^j \to x \in \mathcal{H}$, then $x \in M^\perp$. If $\overline{M}$ denotes the smallest closed subspace that contains M, then we have from the projection lemma that

$$\mathcal{H} = \overline{M} \oplus M^\perp. \tag{1}$$

This notation, $\oplus$ (called the **orthogonal sum**), means that for every $x \in \mathcal{H}$ there exist $y_1 \in \overline{M}$ and $y_2 \in M^\perp$ such that $x = y_1 + y_2$. Obviously, y_1 and y_2 are unique. y_2 is called the **normal vector** to M through x. The geometric intuition behind (1) is that if $x \in \mathcal{H}$ and M is a closed subspace, then the best least squares fit to x in M is given by $x - y_2$.

To prove (1), pick any $x \in \mathcal{H}$ and consider $\mathcal{C} = \{z \in \mathcal{H} : \ z = x - y, y \in \overline{M}\}$. Clearly, $\mathcal{C}$ is a closed convex set and hence there is $z_0 \in \mathcal{C}$ such that $\|z_0\| = \inf\{\|z\| : z \in \mathcal{C}\}$. Similar to the proof in Sect. 2.8, we find that z_0 is orthogonal to $\overline{M}$, $y_0 := x - z_0 \in \overline{M}$ and thus (1) is proved. It is easy to see that $\overline{M}^\perp = M^\perp$.

The reader is invited to prove the principle of uniform boundedness. That is, whenever $\{l^i\}$ is a collection of bounded linear functionals on $\mathcal{H}$ such that for every $x \in \mathcal{H}$ $\sup_i |l^i(x)| < \infty$, then $\sup_i \|l^i\| < \infty$.

Up to this point our comments concerned analogies with L^p-spaces; with the exception of (1), Hilbert-spaces have not seemed to be much different from L^p-spaces. The essential differences will be discussed next.

An orthonormal basis is a key notion in Euclidean spaces (which themselves are special examples of Hilbert-spaces) and this can be carried over to all Hilbert-spaces. Call a set $\mathcal{S} = \{w_1, w_2, \dots\}$ of vectors in $\mathcal{H}$ an **orthonormal set** if $(w_i, w_j) = \delta_{i,j}$ for all $w_i, w_j \in \mathcal{S}$. Here $\delta_{i,j} = 1$ if $i = j$

and $\delta_{i,j} = 0$ if $i \neq j$. If $x \in \mathcal{H}$ is given, one may ask for the best quadratic fit to x by linear combinations of vectors in $\mathcal{S}$. If $\mathcal{S}$ is a finite set, then the answer is $x_N = \sum_{j=1}^{N} (w_j, x) w_j$ as is easily shown. Clearly,

$$0 \leq \|x - x_N\|^2 = \|x\|^2 - 2\,\mathrm{Re}(x, x_N) + \|x_N\|^2 = \|x\|^2 - \sum_{j=1}^{N} |(w_j, x)|^2$$

and we obtain the important **inequality of Bessel**

$$\sum_{j=1}^{N} |(w_j, x)|^2 \leq \|x\|^2.$$

From now on we shall assume that $\mathcal{H}$ is a **separable Hilbert-space**, i.e., there exists a countable, dense set $\mathcal{C} = \{u_1, u_2, \dots\} \subset \mathcal{H}$. (Nonseparable Hilbert-spaces are unpleasant, used rarely and best avoided.) Thus, for every element $x \in \mathcal{H}$ and for $\varepsilon > 0$, there exists N such that $\|x - u_N\| < \varepsilon$. From $\mathcal{C}$ we can construct a countable set $\mathcal{B} = \{w_1, w_2, \dots\}$ as follows. Define $w_1 := u_1/\|u_1\|$, and then recursively define $w_k := v_k/\|v_k\|$, where

$$v_k := u_k - \sum_{j=1}^{k-1} (w_j, u_k) w_j.$$

If $v_k = 0$, then throw out u_k from $\mathcal{C}$ and continue on. The set $\mathcal{B}$ is easily seen to be orthonormal and this constructive procedure for obtaining orthonormal sets is called the **Gram–Schmidt procedure**.

Suppose there is an $x \in \mathcal{H}$ such that $(x, w_k) = 0$ for all k. We claim that then $x = 0$. Recalling that $\mathcal{C} \subset \mathcal{H}$ is dense, pick $\varepsilon > 0$ and then find $u_N \in \mathcal{C}$ such that $\|x - u_N\| < \varepsilon$. By the Gram–Schmidt procedure we know that

$$u_N = v_N + \sum_{j=1}^{N-1} (w_j, u_N) w_j \quad \text{for any } N.$$

Since v_N is proportional to w_N, the condition $(x, w_k) = 0$ for all k implies that $(x, u_N) = 0$. Since $\varepsilon^2 > \|x - u_N\|^2 = \|x\|^2 + \|u_N\|^2$, we find that $\|x\| < \varepsilon$. But ε is arbitrary, so $x = 0$, as claimed.

By Bessel's inequality, the sequence

$$x_M := \sum_{j=1}^{M} (w_j, x) w_j$$

is a Cauchy sequence and hence there is an element $y \in \mathcal{H}$ such that $\|y - x_M\| \to 0$ as $M \to \infty$. Clearly, $(x - y, w_j) = 0$ for all j, and hence $x = y$. Thus we have arrived at the important fact that the set $\mathcal{B}$ is an **orthonormal basis** for our Hilbert-space, i.e., every element $x \in \mathcal{H}$ can be expanded as a **Fourier series**

$$x = \sum_{j=1}^{D} (w_j, x) w_j, \tag{2}$$

where D, the **dimension** of $\mathcal{H}$, is finite or infinite (we shall always write ∞ for brevity). The numbers (w_j, x) are called the **Fourier coefficients** of the element x (with respect to the basis $\mathcal{B}$, of course). It is important to note that

$$\sum_{j=1}^{\infty} (w_j, x) w_j$$

stands for the limit of the sequence

$$x_M = \sum_{j=1}^{M} (w_j, x) w_j$$

in $\mathcal{H}$ as $M \to \infty$.

It is now very simple to show the analogue of Theorem 2.18, that every ball in a separable Hilbert-space is weakly sequentially compact. To be precise, let x_i be a bounded sequence in $\mathcal{H}$. Then there exists a subsequence x_{i_k} and a point $x \in \mathcal{H}$ such that

$$\lim_{k \to \infty} (x_k, y) = (x, y)$$

for every $y \in \mathcal{H}$. Again, we leave the easy details to the reader.

There are many more fundamental points to be made about Hilbert-spaces, such as linear operators, self-adjoint operators and the spectral theorem. All these notions are not only fairly deep mathematically, but they are also the key to the interpretation of quantum mechanics; indeed, many concepts in Hilbert-space theory were developed under the stimulus of quantum mechanics in the first half of the twentieth century. There are many excellent texts that cover these topics.

Exercises for Chapter 2

1. Show that for any two nonnegative numbers a and b

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

where $1 \leq p, q \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Use this to give another proof of Theorem 2.3 (Hölder's inequality).

2. Prove 2.1(6) and the statement that when $\infty \geq r \geq q \geq 1$, $f \in L^r(\Omega) \cap L^q(\Omega) \Rightarrow f \in L^p(\Omega)$ for all $r \geq p \geq q$.

3. [Banach-Saks] proved that after passing to a *subsequence* the c_k^j in Theorem 2.13 can be taken to be $c_k^j = 1/j$. Prove this for $L^2(\Omega)$, i.e., for Hilbert spaces.

4. The penultimate sentence in the remark in Sect. 2.5 is really a statement about nonnegative numbers. Prove it, i.e., for $1 \leq p \leq 2$ and for $0 < b < a$

$$(a+b)^p + (a-b)^p \geq 2a^p + p(p-1)a^{p-2}b^2.$$

5. Referring to Theorem 2.5, assume that $1 < p \leq 2$ and that f and g lie on the unit sphere in L^p, i.e., $\|f\|_p = \|g\|_p = 1$. Assume also that $\|f - g\|_p$ is small. Draw a picture of this situation. Then, using Exercise 4, explain why 2.5(2) shows that the unit sphere is 'uniformly convex'. Explain also why 2.5(1) shows that the unit sphere is 'uniformly smooth', i.e., it has no corners.

6. As needed in the proof of Theorem 2.13 (strongly convergent convex combinations), prove that 'Cauchy sequences of Cauchy sequences are Cauchy sequences'. (In particular, state clearly what this means.)

7. Assume that f and g are in $L^1(\mathbb{R}^n)$. Prove that the convolution $f * g$ in 2.15(1) is a measurable function and that this function is in $L^1(\mathbb{R}^n)$.

8. Prove that a strongly convergent sequence in $L^p(\mathbb{R}^n)$ is also a Cauchy sequence.

9. In Sect. 2.9 three ways are shown for which an $L^p(\mathbb{R}^n)$ sequence f^k can converge weakly to zero but f^k does not convergence to anything strongly. Verify this for the three examples given in 2.9.

10. Let f be a real-valued, measurable function on $\mathbb{R}$ that satisfies the equation

$$f(x+y) = f(x) + f(y)$$

for all x, y in $\mathbb{R}$. Prove that $f(x) = Ax$ for some number A.

► *Hint.* Prove this when f is continuous by examining f on the rationals. Next, convolve $\exp[if(x)]$ with a j_ε of compact support. The convolution is continuous!

11. With the usual $j_\varepsilon \in C_c^\infty$, show that if f is continuous then $j_\varepsilon * f(x)$ converges to $f(x)$ for all x, and it does so uniformly on each compact subset of $\mathbb{R}^n$.

12. Deduce Schwarz's inequality $|(x,y)| \leq \sqrt{(x,x)}\sqrt{(y,y)}$ from 2.21 (i)–(iv) *alone*. Determine all the cases of equality.

13. Prove the analogue of Lemma 2.8 (Projection on convex sets) for Hilbert-spaces.

14. For any (not necessarily closed) subspace M show that $M^\perp$ is closed and that $\overline{M}^\perp = M^\perp$.

15. Prove Riesz's representation theorem, Theorem 2.14, for Hilbert-spaces.

16. Prove the principle of uniform boundedness for Hilbert-spaces by imitating the proof in Sect. 2.12.

17. Prove that every bounded sequence in a separable Hilbert-space has a weakly convergent subsequence.

18. Prove that every convex function has a support plane at every x in the interior of its domain, as claimed in Sect. 2.1. See also Exercise 3.1.

19. Prove 2.9(4).

20. Find a sequence of bounded, measurable sets in $\mathbb{R}$ whose characteristic functions converge weakly in $L^2(\mathbb{R})$ to a function f with the property that $2f$ is a characteristic function. How about the possibility that $f/2$ is a characteristic function?

21. At the end of the proof of Theorem 2.6 (Differentiability of norms) there is a displayed pair of inequalities, valid for $|t| \leq 1$:

$$|f|^p - |f-g|^p \leq \frac{1}{t}\{|f+tg|^p - |f|^p\} \leq |f+g|^p - |f|^p.$$

Write out a complete proof of these two inequalities.

22. Prove the $\boldsymbol{p,q,r}$ **theorem**: *Suppose that* $1 \le p < q < r \le \infty$ *and that* f *is a function in* $L^p(\Omega, \mathrm{d}\mu) \cap L^r(\Omega, \mathrm{d}\mu)$ *with* $\|f\|_p \le C_p < \infty$, $\|f\|_r \le C_r < \infty$, *and* $\|f\|_q \ge C_q > 0$. *Then there are constants* $\varepsilon > 0$ *and* $M > 0$, *depending only on* p, q, r, C_p, C_q, C_r, *such that* $\mu(\{x : |f(x)| > \varepsilon\}) > M$.

    In fact, if we define S, T by $q\, C_p^p S^{q-p} = (q-p)C_q^q/4$ and $q\, C_r^r T^{q-r} = (r-q)C_q^q/4$, then we may take $\varepsilon = S$ and $M = |T^q - S^q|^{-1} C_q/2$. (See [Fröhlich–Lieb–Loss].)

    Show, conversely, that without knowledge of C_q, $\mu(\{x : |f(x)| > \varepsilon\})$ can be arbitrarily small for any fixed number $\varepsilon > 0$.

    ▶ *Hint.* Use the layer cake principle to evaluate the various norms.

23. Find a sequence of functions with the property that f^j converges to 0 in $L^2(\Omega)$ weakly, to 0 in $L^{3/2}(\Omega)$ strongly, but it does not converge to 0 strongly in $L^2(\Omega)$.