

MASDOC A1  
Linear Partial Differential Equations

James C. Robinson



---

# Contents

<b>1</b>	<b>Introduction: Hilbert spaces and PDEs</b>	<i>page</i> 1
1.1	Inner products, norms, and Hilbert spaces	1
1.2	Closest points in closed linear subspaces	3
1.3	Linear operators and dual spaces	5
1.4	The Riesz Representation Theorem	6
1.5	Poisson's equation	9
1.6	'Variational formulation'	12
1.7	Non-symmetric $B$ : the Lax–Milgram Lemma	13
<b>2</b>	<b>Function spaces</b>	18
2.1	Euclidean spaces	18
2.2	The sequence spaces $\ell^p$	18
2.3	Density and separability	21
2.4	Spaces of continuous and differentiable functions	22
2.4.1	Spaces of continuous functions	22
2.4.2	Multi-index notation and spaces of differentiable functions	24

2.4.3	'Test functions'	24
2.4.4	Mollification	25
2.5	The Lebesgue integral on $\mathbb{R}$	26
2.6	The Lebesgue spaces $L^p$	29
2.7	Locally integrable functions	34
2.8	Sobolev spaces	36
2.8.1	Weak derivatives	36
2.8.2	Sobolev spaces	38
2.8.3	Completions	39
2.8.4	Density results for $H^k(\Omega)$	41
2.8.5	Extension theorems	44
2.9	The Fourier transform and Sobolev spaces	48
2.9.1	Compact embeddings	51
2.9.2	Sobolev spaces and $L^p$ spaces	53
2.9.3	Scaling and inequalities	54
2.9.4	Boundary values and the trace operator	55
<b>3</b>	<b>Elliptic PDEs</b>	<b>57</b>
3.1	Elliptic regularity	60
3.1.1	Interior regularity	60
3.1.2	Boundary regularity	64
<b>4</b>	<b>Eigenvalues, eigenfunctions, and bases in Hilbert spaces</b>	<b>66</b>
4.1	Eigenvalues and the spectrum	66
4.2	The spectrum of a compact symmetric operator	67

4.3	Bases in Hilbert spaces	71
4.4	Eigenfunctions as basis elements; the Hilbert–Schmidt Theorem	74
4.5	Eigenvalues and eigenfunctions for elliptic PDEs	76
<b>5</b>	<b>Linear parabolic problems</b>	<b>78</b>
5.1	Banach-space valued function spaces and ‘Sobolev-type’ results	79
5.2	Weak solutions of parabolic problems	83
5.3	Galerkin approximations	84
<b>6</b>	<b>Nonlinear equations, weak convergence</b>	<b>91</b>
6.1	Dual spaces, the Hahn–Banach Theorem, and weak convergence	91
6.2	Weak compactness and our linear parabolic PDE	96



# 1

---

## Introduction: Hilbert spaces and PDEs

To begin with, we outline a ‘simple’ approach to proving the existence and uniqueness of solutions of Poisson’s equation with Dirichlet boundary conditions:

$$-\Delta u = f \quad \text{with} \quad u|_{\partial\Omega} = 0, \quad (1.1)$$

where  $\Omega \subset \mathbb{R}^n$  is a smooth bounded domain. Our approach uses ideas from the theory of Hilbert spaces, so we recall some of these ideas.

A normed vector space<sup>1</sup> is *complete* if every Cauchy sequence converges<sup>2</sup>.

A complete normed space is called a *Banach space*.

### 1.1 Inner products, norms, and Hilbert spaces

Examples of Banach spaces:  $\mathbb{R}^n$ ,  $\ell^p$ ,  $L^p(\Omega)$ ,  $C^0(\Omega)$ ,  $C^0(\bar{\Omega})$ ,  $C^{0,\alpha}(\Omega)$ ,  $W^{k,p}(\Omega)$ ,  $H^k(\Omega)$ ...

An *inner product* on a normed vector space is a map  $(\cdot, \cdot) : X \times X \rightarrow \mathbb{K}$  such that

- (i)  $(x, x) \geq 0$  for all  $x \in X$ , with equality iff  $x = 0$ ;

<sup>1</sup> All vector spaces will be over  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ , and usually over  $\mathbb{R}$ .

<sup>2</sup> • A *norm* on a vector space  $X$  is a map  $\|\cdot\| : X \rightarrow [0, \infty)$  such that (i)  $\|x\| = 0$  iff  $x = 0$ ; (ii)  $\|\lambda x\| = |\lambda| \|x\|$  for every  $\lambda \in \mathbb{K}$  and  $x \in X$ ; and (iii)  $\|x + y\| \leq \|x\| + \|y\|$  for all  $x, y \in X$ .

• A sequence  $\{x_n\} \in X$  is *Cauchy* if for every  $\epsilon > 0$  there exists an  $N$  such that  $\|x_n - x_m\| < \epsilon$  for all  $n, m \geq N$ .

• A sequence  $\{x_n\}$  in  $X$  *converges* to  $x \in X$  if  $\|x_n - x\| \rightarrow 0$  as  $n \rightarrow \infty$ .

- (ii)  $(x + y, z) = (x, z) + (y, z)$  for all  $x, y, z \in X$ ;
- (iii)  $(\alpha x, y) = \alpha(x, y)$  for all  $\alpha \in \mathbb{K}$ ,  $x, y \in X$ ; and
- (iv)  $(x, y) = \overline{(y, x)}$ .

An inner product induces a norm via

$$\|x\|^2 = (x, x). \quad (1.2)$$

**Lemma 1.1 (Cauchy-Schwarz inequality)** *If  $(\cdot, \cdot)$  is an inner product and  $\|\cdot\|$  is defined as in (1.2) then*

$$|(x, y)| \leq \|x\| \|y\| \quad \text{for all } x, y \in X.$$

*In particular,  $\|\cdot\|$  defines a norm on  $X$ .*

*Proof* Consider

$$\|x - \lambda y\|^2 = (x - \lambda y, x - \lambda y) = \|x\|^2 - \lambda(y, x) - \bar{\lambda}(x, y) + |\lambda|^2 \|y\|^2 \geq 0;$$

setting  $\lambda = (x, y)/\|y\|^2$  yields

$$\|x\|^2 - 2 \frac{|(x, y)|^2}{\|y\|^2} + \frac{|(x, y)|^2}{\|y\|^2} \geq 0$$

and the inequality holds.

To show that  $\|\cdot\|$  is a norm, properties (i) and (ii) are clear. To show (iii), note that

$$\begin{aligned} \|x + y\|^2 &= (x + y, x + y) = \|x\|^2 + (x, y) + (y, x) + \|y\|^2 \\ &\leq \|x\|^2 + 2|(x, y)| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \\ &\leq (\|x\| + \|y\|)^2. \end{aligned}$$

□

A complete inner product space is called a *Hilbert space*.

Every Hilbert space is a Banach space. The converse is not true, since in a Hilbert space  $H$  the norm must satisfy the parallelogram law<sup>1</sup>

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \text{for all } x, y \in H. \quad (1.3)$$

<sup>1</sup> In fact this characterises those norms that can be derived from an inner product. If a norm comes from an inner product you can recover the inner product from the norm using the ‘polarisation identity’  $4(x, y) = \|x + y\|^2 - \|x - y\|^2$  in the real case (the complex case is messier; in both cases you prove the identity by expanding the norms as inner products). If a norm on a real vector space satisfies the parallelogram law and you define a map  $(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$  using this polarisation identity, you will end up with an inner product (this is quite awkward, if not difficult, to prove).

(To prove simply expand both sides using the inner product.)

## 1.2 Closest points in closed linear subspaces

We now use the parallelogram law to show that in a Hilbert space, there is always a ‘closest point’ in a closed linear subspace<sup>1</sup>.

**Theorem 1.2** *Let  $U$  be a closed linear subspace<sup>2</sup> of a Hilbert space  $H$  and let  $x \in H$ . Then there exists a unique  $\hat{a} \in U$  such that*

$$\|x - \hat{a}\| = \inf\{\|x - a\| : a \in U\}.$$

Furthermore,  $x - \hat{a} \in U^\perp$ , where

$$U^\perp = \{u \in H : (u, x) = 0 \quad \text{for all } x \in U\}.$$

The map  $P : H \rightarrow U$  given by  $x \mapsto \hat{a}$  is the orthogonal projection of  $x$  onto  $U$ ;  $P^2 = P$  and  $\|Px\| \leq \|x\|$ .

Two elements  $x$  and  $y$  of an inner product space are said to be orthogonal if  $(x, y) = 0$ . (We sometimes write  $x \perp y$ .)  $U^\perp$  is the *orthogonal complement* of  $U$ .

*Proof* Set  $\delta = \inf\{\|x - a\| : a \in A\}$  and find a sequence  $a_n \in A$  such that

$$\|x - a_n\|^2 \leq \delta^2 + \frac{1}{n}. \quad (1.4)$$

We will show that  $\{a_n\}$  is a Cauchy sequence. To this end, we use the parallelogram law:

$$\|(x - a_n) + (x - a_m)\|^2 + \|(x - a_n) - (x - a_m)\|^2 = 2[\|x - a_n\|^2 + \|x - a_m\|^2].$$

Which gives

$$\|2x - (a_n + a_m)\|^2 + \|a_n - a_m\|^2 < 4\delta^2 + \frac{2}{m} + \frac{2}{n}$$

or

$$\|a_n - a_m\|^2 \leq 4\delta^2 + \frac{2}{m} + \frac{2}{n} - 4\|x - \frac{1}{2}(a_n + a_m)\|^2.$$

<sup>1</sup> Linear subspaces of infinite-dimensional spaces are not always closed. Can you find an example? However, for any set  $U$ ,  $U^\perp$  is always a closed linear subspace: clearly linear, and closed since if  $x_n \rightarrow x$  with  $x_n \in U^\perp$ , then  $(x, u) = \lim_{n \rightarrow \infty} (x_n, u) = 0$  for every  $u \in U$ .

<sup>2</sup> There also exists a closest point in any closed convex subset of  $H$ : a set  $U$  is convex if for every  $u, v \in U$ ,  $\lambda u + (1 - \lambda)v \in U$  for all  $\lambda \in [0, 1]$ .

Since  $A$  is convex,  $a_n + a_m \in A$ , and so  $\|x - \frac{1}{2}(a_n + a_m)\|^2 \geq \delta^2$ , which gives

$$\|a_n - a_m\|^2 \leq \frac{2}{m} + \frac{2}{n}.$$

It follows that  $\{a_n\}$  is Cauchy, and so  $a_n \rightarrow \hat{a}$ . Since  $A$  is closed,  $\hat{a} \in A$ .

To show that  $\hat{a}$  is unique, suppose that  $\|u - a^*\| = \delta$  with  $a^* \neq \hat{a}$ . Then  $\|u - \frac{1}{2}(a^* + \hat{a})\| \geq \delta$  since  $A$  is a linear subspace, and so, using the parallelogram law again,

$$\|a^* - \hat{a}\|^2 \leq 4\delta^2 - 4\delta^2 = 0,$$

i.e.  $a^* = \hat{a}$  and  $\hat{a}$  is unique.

Now consider  $v = x - \hat{a}$ ; the claim is that  $v \in U^\perp$ , i.e. that

$$(v, y) = 0 \quad \text{for all } y \in U.$$

Consider  $\|x - (\hat{a} - ty)\| = \|v + ty\|$ ; then

$$\begin{aligned} \Delta(t) &= \|v + ty\|^2 = (v + ty, v + ty) \\ &= \|v\|^2 + (ty, v) + (v, ty) + \|y\|^2 \\ &= \|v\|^2 + t(y, v) + \overline{t(y, v)} + |t|^2\|y\|^2 \\ &= \|v\|^2 + 2\operatorname{Re}\{t(y, v)\} + |t|^2\|y\|^2. \end{aligned}$$

We know from the construction of  $\hat{a}$  that  $\|v + ty\|$  is minimal when  $t = 0$ . If  $t$  is real then this implies that  $d\Delta/dt(0) = 2\operatorname{Re}\{(y, v)\} = 0$ . If  $t = is$ , with  $s$  real, then  $d\Delta(is)/ds = -2\operatorname{Im}\{(y, v)\} = 0$ . So  $(y, v) = 0$  for any  $y \in U$ , i.e.  $v \in U^\perp$  is claimed.

Clearly  $P^2 = P$ , and since

$$\|x\|^2 = \|\hat{a}\|^2 + 2(\hat{a}, x - \hat{a}) + \|x - \hat{a}\|^2,$$

the inequality  $\|Pu\| \leq \|u\|$  is clear.  $\square$

Note that when  $U$  is a one-dimensional subspace<sup>1</sup>, the closest point can be given explicitly: take any  $u \in U$  with  $\|u\| = 1$ ; then for the closest point to be  $tu$  we need

$$(x - tu, u) = 0 \quad \Rightarrow \quad (x, u) - t\|u\|^2 = 0 \quad \Rightarrow \quad t = (x, u), \quad (1.5)$$

i.e. the closest point is  $(x, u)u$ .

<sup>1</sup> This approach can be extended to any space spanned by a finite/countable collection of orthonormal vectors  $\{e_j\}$ , with the closest point being given by  $\sum(x, e_j)e_j$ .

### 1.3 Linear operators and dual spaces

Much of functional analysis concerns operators - particularly linear operators - between different spaces. An operator  $L : X \rightarrow Y$  is *linear* if

$$L(\alpha x_1 + \beta x_2) = \alpha Lx_1 + \beta Lx_2 \quad \text{for all} \quad \alpha, \beta \in \mathbb{K}, \quad x_1, x_2 \in X.$$

A linear operator  $L : X \rightarrow Y$  is *bounded* if there exists a constant  $M$  such that

$$\|Lx\|_Y \leq M\|x\|_X; \quad (1.6)$$

we denote the space of all bounded linear maps<sup>1</sup> from  $X$  into  $Y$  by  $\mathcal{L}(X, Y)$ , and define the *operator norm* of  $L$  (as an operator from  $X$  into  $Y$ ) by

$$\|L\|_{\mathcal{L}(X, Y)} = \inf\{M : \|Lx\|_Y \leq M\|x\|_X\}. \quad (1.7)$$

(When there is no ambiguity this is often abbreviated to  $\|L\|$ .)

The space  $\mathcal{L}(X, Y)$  is a vector space (with the obvious definitions of addition and scalar multiplication), and  $\|\cdot\|_{\mathcal{L}(X, Y)}$  is a norm on this space. What is somewhat remarkable is that if  $Y$  is complete then so is  $\mathcal{L}(X, Y)$ .

**Theorem 1.3** *If  $X$  is a normed space and  $Y$  is a Banach space, then  $\mathcal{L}(X, Y)$  is complete (i.e. is a Banach space).*

*Proof* Let  $\{L_n\}$  be a Cauchy sequence in  $\mathcal{L}(X, Y)$ . We identify a candidate  $L$  for the limit of the sequence  $\{L_n\}$ , then show that this is an element of  $\mathcal{L}(X, Y)$ , and that  $L_n \rightarrow L$  (wrt the operator norm).

Since  $\{L_n\}$  is Cauchy, for any  $\epsilon > 0$  there exists an  $N$  such that

$$\|L_n - L_m\|_{\mathcal{L}(X, Y)} < \epsilon \quad \text{for all} \quad n, m \geq N.$$

It follows that for any fixed  $x \in X$ ,

$$\|L_n x - L_m x\|_Y = \|(L_n - L_m)x\|_Y < \epsilon\|x\|, \quad (1.8)$$

i.e. that  $\{L_n x\}$  is a Cauchy sequence in  $Y$ . So for each  $x \in X$  we can define

$$Lx = \lim_{n \rightarrow \infty} L_n x.$$

We claim that (i)  $L \in \mathcal{L}(X, Y)$  and (ii) that  $L_n \rightarrow L$  (wrt  $\|\cdot\|_{\mathcal{L}(X, Y)}$ ).

<sup>1</sup> There are many unbounded linear maps; in particular, differential operators...

To show (i), we first check that  $L$  is linear:

$$\begin{aligned} L(\alpha x_1 + \beta x_2) &= \lim_{n \rightarrow \infty} L_n(\alpha x_1 + \beta x_2) \\ &= \lim_{n \rightarrow \infty} [\alpha L_n x_1 + \beta L_n x_2] \\ &= \alpha \lim_{n \rightarrow \infty} L_n x_1 + \beta \lim_{n \rightarrow \infty} L_n x_2 \\ &= \alpha L x_1 + \beta L x_2, \end{aligned}$$

and that  $L$  is bounded: take  $m \rightarrow \infty$  in (1.8) to obtain

$$\|L_n x - Lx\|_Y < \epsilon \|x\|.$$

This shows that  $L = (L - L_n) + L_n$  is bounded, and that  $L_n$  converges to  $L$ .  $\square$

The dual space of a normed space  $X$  (over  $\mathbb{K}$ ) is the space  $X^* = \mathcal{L}(X, \mathbb{K})$ . We investigate the dual space of Hilbert space in the next section. When not ambiguous for  $f \in X^*$  we write

$$\|f\|_* = \|f\|_{X^*} = \|f\|_{\mathcal{L}(X, \mathbb{K})}.$$

#### 1.4 The Riesz Representation Theorem

First we show that if  $H$  is a Hilbert space, from any  $y \in H$  we can construct an element of  $H^*$  with the same norm, using the inner product.

**Example 1.4** *Let  $U$  be a Hilbert space. Given any  $y \in H$ , define*

$$l_y(x) = (x, y). \tag{1.9}$$

*Then  $l_y$  is clearly linear, and*

$$|l_y(x)| = |(x, y)| \leq \|x\| \|y\|$$

*using the Cauchy-Schwarz inequality. It follows that  $l_y \in H^*$  with  $\|l_y\| \leq \|y\|$ . Choosing  $x = y$  in (1.9) shows that*

$$|l_y(y)| = (y, y) = \|y\|^2$$

*and hence  $\|l_y\| = \|y\|$ .*

The Riesz Representation Theorem shows that this example can be ‘reversed’, i.e. every linear functional on  $H$  corresponds to some inner product:

**Theorem 1.5 (Riesz Representation Theorem)** *For every bounded linear functional  $f$  on a Hilbert space  $H$  there exists a unique element  $y \in H$  such that*

$$f(x) = (x, y) \quad \text{for all } x \in H; \quad (1.10)$$

furthermore  $\|y\|_H = \|f\|_{H^*}$ .

*Proof* Let

$$K = \text{Ker}(f) = \{x \in H : f(x) = 0\},$$

which is a closed linear subspace of  $H$ . [Clearly a linear subspace; closed since if  $x_n \in \text{Ker}(f)$  and  $x_n \rightarrow x$  then

$$|f(x)| = |f(x) - f(x_n)| = |f(x - x_n)| \leq \|f\|_* \|x - x_n\|;$$

this holds for all  $n$ , hence  $|f(x)| = 0$ .]

If  $f = 0$  then  $K = H$  and we set  $y = 0$ .

Otherwise, if  $f \neq 0$  then  $K \neq H$ ; it follows from Theorem 1.2 that  $K^\perp$  is non-empty. In fact,  $K^\perp$  is a one-dimensional linear subspace of  $H$ . Indeed, given  $u, v \in K^\perp$  we have

$$f(f(u)v - f(v)u) = 0. \quad (1.11)$$

Since  $u, v \in K^\perp$  it follows that  $f(u)v - f(v)u \in K^\perp$ , while (1.11) shows that  $f(u)v - f(v)u \in K$ . It follows<sup>1</sup> that

$$l(u)v - l(v)u = 0,$$

and so  $u$  and  $v$  are proportional.

Therefore we can choose  $z \in K$  such that  $\|z\| = 1$ , and decompose any  $x \in H$  as<sup>2</sup>

$$x = (x, z)z + w \quad \text{with } w \in K.$$

(We are using Theorem 1.2 along with the observation in 1.5.) Therefore

$$l(x) = (x, z)l(z) = (x, \overline{l(z)}z).$$

Setting  $y = \overline{l(z)}z$  we obtain (1.10).

To show that this choice of  $y$  is unique, suppose that

$$(x, y) = (x, \hat{y}) \quad \text{for all } x \in H.$$

Then  $(x, y - \hat{y}) = 0$  for all  $x \in H$ , i.e.  $y - \hat{y} \in H^\perp = \{0\}$ .

<sup>1</sup> We always have  $U \cap U^\perp = \{0\}$ : if  $x \in U$  and  $x \in U^\perp$  then  $\|x\|^2 = (x, x) = 0$ .

<sup>2</sup> If  $U$  is a closed linear subspace of  $H$  then  $(U^\perp)^\perp = U$ , which we are in fact using here...

Finally, the calculation in (1.4) shows the equality of the norms of  $y$  and  $f$ .  $\square$

The map  $f \mapsto y$  is known as the *Riesz mapping*. Note that it is a linear isometry (this just means that  $\|f\|_* = \|y\|$ ); as is its inverse. It follows that  $H$  and  $H^*$  are (isometrically) isomorphic (which we write as  $H \simeq H^*$ ).

The following useful corollary is almost immediate. We say that a map  $B : H \times H \rightarrow \mathbb{R}$  is a *bilinear form* if  $B(u, v)$  is linear in both  $u$  and  $v$ , i.e.

$$B(\alpha u_1 + \beta u_2, v) = \alpha B(u_1, v) + \beta B(u_2, v) \quad \text{for all } \alpha, \beta \in \mathbb{R}, u_1, u_2, v \in H$$

and

$$B(u, \alpha v_1 + \beta v_2) = \alpha B(u, v_1) + \beta B(u, v_2) \quad \text{for all } \alpha, \beta \in \mathbb{R}, u, v_1, v_2 \in H.$$

**Corollary 1.6** *Suppose that  $B : H \times H \rightarrow \mathbb{R}$  is a bilinear form that is*

(i) *bounded, i.e. there exists an  $\alpha \geq 0$  such that*

$$|B(u, v)| \leq \alpha \|u\| \|v\| \quad \text{for all } u, v \in H,$$

(ii) *coercive, i.e. there exists a  $\beta > 0$  such that*

$$B(u, u) \geq \beta \|u\|^2 \quad \text{for all } u \in H,$$

and

(iii) *symmetric,*

$$B(u, v) = B(v, u) \quad \text{for all } u, v \in H.$$

Then for any  $f \in H^*$  there exists a unique  $u \in H$  such that

$$B(u, v) = f(v) \quad \text{for all } v \in H, \tag{1.12}$$

and

$$\|u\| \leq \frac{\|f\|_*}{\beta}. \tag{1.13}$$

*Proof*  $B$  is an alternative inner product on  $H$ ; from (ii)  $B(u, u) \geq 0$  with equality iff  $u = 0$ ;  $B$  satisfies the linearity requirements of an inner product; and  $B$  is symmetric by assumption. We can immediately apply the Riesz Representation Theorem to deduce that existence of a unique  $u$  satisfying (1.12). The bound on  $u$  in (1.13) follows immediately from (ii), since

$$\beta \|u\|^2 \leq B(u, u) = f(u) \leq \|f\|_* \|u\|.$$

$\square$

In a theorem below (the Lax–Milgram Lemma) we will remove the requirement that  $B$  is symmetric.

### 1.5 Poisson's equation

We return to Poisson's equation

$$-\Delta u = f \quad x \in \Omega, \quad u|_{\partial\Omega} = 0.$$

Multiply this equation by a  $C^\infty$  function  $\varphi$  that is zero on the boundary, and integrate over  $\Omega$  to obtain

$$-\int_{\Omega} \varphi(x) \Delta u(x) \, dx = \int_{\Omega} f(x) \varphi(x) \, dx.$$

Now integrate the first term by parts<sup>1</sup> to obtain

$$\int_{\Omega} \nabla \varphi \cdot \nabla u \, dx = \int_{\Omega} f \varphi \, dx. \quad (1.14)$$

We want to recast this in a form suitable to apply Corollary 1.6. To do this we require an appropriate Hilbert space.

Let

$$\tilde{L}^2(\Omega) = \{f \in C^\infty(\Omega) : \int_{\Omega} |f(x)|^2 \, dx < \infty\};$$

we define the  $L^2(\Omega)$  norm of  $f$  to be

$$\|f\|_{L^2}^2 = \left( \int_{\Omega} |f(x)|^2 \, dx \right)^{1/2}; \quad (1.15)$$

this is induced by the  $L^2$ -inner product

$$(f, g)_{L^2} = \int_{\Omega} f(x)g(x) \, dx. \quad (1.16)$$

It is straightforward to check that  $(\cdot, \cdot)_{L^2}$  is indeed an inner product (and

<sup>1</sup> This is the divergence theorem:

$$\int_{\Omega} \nabla \cdot (\varphi \nabla u) \, dx = \int_{\partial\Omega} \varphi \nabla u \cdot n \, dS;$$

the right-hand side is zero since  $\varphi = 0$  on  $\partial\Omega$ , and the left-hand side is

$$\int_{\Omega} \nabla \varphi \cdot \nabla u \, dx + \int_{\Omega} \varphi \Delta u \, dx.$$

hence that  $\|\cdot\|_{L^2}$  is indeed a norm). The space  $\tilde{L}^2(\Omega)$  is in fact not complete with respect to this norm (exercise: find a sequence in  $\tilde{L}^2(\Omega)$  that is Cauchy with respect to this norm but does not have a limit in  $\tilde{L}^2(\Omega)$ ); its completion is the Lebesgue space  $L^2(\Omega)$  which we will study in some detail later.

We also let

$$\tilde{H}_0^1(\Omega) = \{f \in C^\infty(\Omega) : f|_{\partial\Omega} = 0, \int_{\Omega} |f(x)|^2 + |\nabla f(x)|^2 dx < \infty\},$$

and define the  $H^1$  norm of  $f$  to be given by

$$\|f\|_{H^1}^2 = \int_{\Omega} |f(x)|^2 + |\nabla f(x)|^2 dx = \|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2;$$

this can be induced by the inner product

$$((f, g))_{H^1} = (f, g)_{L^2} + \sum_{j=1}^n (\partial_j f, \partial_j g)_{L^2} =: (f, g)_{L^2} + (\nabla f, \nabla g)_{L^2}.$$

Again,  $\tilde{H}_0^1(\Omega)$  is not in fact complete with respect to this norm; its completion is the Sobolev space  $H_0^1(\Omega)$ .

We can now rewrite (1.14) as

$$(\nabla u, \nabla \varphi)_{L^2} = (f, \varphi)_{L^2}.$$

We want to consider this as an equation in the Hilbert space  $H_0^1$  to which we can apply Corollary 1.6.

First we note that  $(f, \varphi)_{L^2}$  maps  $H_0^1 \subset L^2$  into  $\mathbb{R}$ , and is clearly linear in  $\varphi$ ; it is also a bounded mapping, since

$$|(f, \varphi)_{L^2}| \leq \|f\|_{L^2} \|\varphi\|_{L^2} \leq \|f\|_{L^2} \|\varphi\|_{H^1},$$

using the Cauchy-Schwarz inequality in  $L^2$  and the definition of the  $H^1$  norm.

Now, if we define

$$B(u, v) = (\nabla u, \nabla v)_{L^2}$$

for  $u, v \in H_0^1$ , then this is clearly a symmetric bilinear form. It is bounded since

$$|B(u, v)| = |(\nabla u, \nabla v)_{L^2}| \leq \|\nabla u\|_{L^2} \|\nabla v\|_{L^2} \leq \|u\|_{H^1} \|v\|_{H^1},$$

again using the Cauchy-Schwarz inequality in  $L^2$  and the definition of the  $H^1$  norm.

However, the coercivity is more delicate. For this we require the following *Poincaré's inequality*:

**Lemma 1.7 (Poincaré's inequality)** *Suppose that  $\Omega$  is bounded in one direction, i.e.  $|x_1| \leq d < \infty$ . Then for every  $f \in \tilde{H}_0^1$*

$$\|f\|_{L^2} \leq 2d \|\nabla f\|_{L^2}. \quad (1.17)$$

(The result actually holds in  $H_0^1(\Omega)$  with essentially the same proof, as we will see; note that it follows from the proof that in fact we do not need all the derivatives on the right-hand side, only the one derivative in the  $x_1$  direction.)

*Proof* The inequality is trivial if  $f = 0$ , so we assume that  $f \neq 0$ , and hence that  $\|f\|_{L^2} \neq 0$ . We integrate by parts in the  $x_1$  variable:

$$\|f\|_{L^2}^2 = \int_{\Omega} 1 \cdot |f(x)|^2 dx = - \int_{\Omega} x_1 \frac{\partial}{\partial x_1} |f(x)|^2 + \int_{\partial\Omega} x_1 |u(x)|^2 n_1 dS,$$

where  $n_1$  is the component of the outward normal in the  $x_1$  direction. Using the fact that  $u = 0$  on  $\partial\Omega$  the boundary term vanishes, and hence

$$\begin{aligned} \|f\|_{L^2}^2 &= -2 \int_{\Omega} x_1 f(x) \frac{\partial}{\partial x_1} f(x) dx \\ &\leq 2 \int_{\Omega} |x_1| |f(x)| \left| \frac{\partial}{\partial x_1} f(x) \right| dx \\ &\leq 2d \int_{\Omega} |f(x)| |\nabla f(x)| dx \\ &\leq 2d \left( \int_{\Omega} |f(x)|^2 dx \right)^{1/2} \left( \int_{\Omega} |\nabla f(x)|^2 dx \right)^{1/2} \\ &= 2d \|f\|_{L^2} \|\nabla f\|_{L^2}, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality (in  $L^2$ ) and the fact that  $|\partial f / \partial x_1| \leq |\nabla f|$ ; (1.17) follows if we divide through by  $\|f\|_{L^2}$ .  $\square$

Two norms  $\|\cdot\|_1$  and  $\|\cdot\|_2$  on a space  $X$  are said to be *equivalent* if there exist constants  $c_2 > c_1 > 0$  such that

$$c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1 \quad \text{for all } x \in X.$$

**Corollary 1.8** *If  $\Omega$  is bounded in one direction then the norms  $\|f\|_{H^1}$  and  $\|\nabla f\|_{L^2}$  are equivalent on  $\tilde{H}_0^1(\Omega)$ .*

*Proof*

$$\|\nabla f\|_{L^2}^2 \leq \underbrace{\|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2}_{\|u\|_{H^1}^2} \leq (4d^2 + 1)\|\nabla f\|^2.$$

□

We have therefore shown that  $B(u, u)$  is coercive on  $\tilde{H}_0^1$ :

$$B(u, u) = \|\nabla u\|_{L^2}^2 \geq \frac{1}{1 + 4d^2} \|u\|_{H^1}^2,$$

and hence all the requirements of Corollary 1.6 are satisfied, from which we can deduce that there is a unique solution  $u \in H_0^1(\Omega)$  of the problem

$$B(u, v) = f(v) \quad \text{for all } v \in H_0^1$$

with  $\|u\|_{H^1} \leq c\|f\|_{L^2}$ .

There are some gaps/problems in this argument, mainly arising from the fact that we haven't defined any function spaces properly:

- (i) We derived the equation

$$B(u, \varphi) = (f, \varphi)$$

under the assumption that everything was smooth; in particular, that  $\varphi$  was a  $C^\infty$  function zero on  $\partial\Omega$ .

- (ii) The spaces  $\tilde{L}^2$  and  $\tilde{H}_0^1$  are not complete, so not Hilbert spaces.  
 (iii) We would really like a classical (at least  $C^2$ ) solution of the problem, but this method only gives us a  $u \in H_0^1$ .

## 1.6 'Variational formulation'

PDEs are closely related to certain problems in the 'calculus of variations' (i.e. minimisation/maximisation problems). Here we will show that the abstract formulation of our PDE,

$$\text{find } u \in H \text{ such that } B(u, v) = f(v) \text{ for all } v \in H, \quad (1.18)$$

is the same as the following minimisation problem:

$$\text{find } u \in H \text{ such that } J(u) \leq J(v) \text{ for all } v \in H, \quad (1.19)$$

where  $J(u) = \frac{1}{2}B(u, u) - f(u)$ .

**Theorem 1.9** *Under the conditions on  $B$  and  $f$  imposed in Corollary 1.6 the two problems (1.18) and (1.19) are equivalent.*

*Proof* First, suppose that we have a solution  $u \in H$  of (1.18), and for an arbitrary  $v \in H$  write

$$\begin{aligned} J(v) &= J(u + (v - u)) \\ &= \frac{1}{2}B(u, u) - f(u) + B(u, v - u) - f(v - u) + \frac{1}{2}B(v - u, v - u) \\ &= J(u) + \underbrace{[B(u, v - u) - f(v - u)]}_{= 0 \text{ since } u \text{ solves (1.18)}} + \frac{1}{2}B(v - u, v - u) \\ &\geq J(u) \end{aligned}$$

since  $B(v - u, v - u) \geq \beta \|v - u\|^2 \geq 0$  using coercivity of  $B$ . So  $u$  minimises  $J$ .

On the other hand, if  $u$  minimises  $J$  then for any  $v \in H$  consider

$$\begin{aligned} J(u + tv) &= \frac{1}{2}B(u + tv, u + tv) - f(u + tv) \\ &= \frac{1}{2}B(u, u) - f(u) + t(B(u, v) - f(v)) + \frac{1}{2}t^2B(v, v). \end{aligned}$$

Since we must have  $J(u + tv) \geq J(u)$  for every  $t$ , it follows that

$$B(u, v) = f(v).$$

Since  $v \in H$  was arbitrary this shows that  $u$  solves (1.18).  $\square$

**Corollary 1.10** *Under the conditions of Corollary 1.6, the minimisation problem (1.19) has a unique solution  $u \in H$ .*

### 1.7 Non-symmetric $B$ : the Lax–Milgram Lemma

We now prove the Lax–Milgram Lemma, a version of Corollary 1.6 valid when  $B$  is not symmetric. In fact the argument is somewhat surprising in the light of where we began, namely a linear PDE, which we could write as

$$Lu = f,$$

where  $L$  is some linear operator. By taking inner products we turned this into an abstract problem involving a bilinear form  $B : H \times H \rightarrow \mathbb{R}$ ,

$$B(u, v) = f(v) \quad \text{for all } v \in H.$$

The main ‘trick’ in the proof of the Lax–Milgram Lemma is to turn this problem back into a linear equation; properties of the bilinear form then guarantee the existence of a solution of this linear equation.

**Theorem 1.11 (Lax–Milgram Lemma)** *Suppose that  $B : H \times H \rightarrow \mathbb{R}$  is a bilinear form that is*

(i) *bounded, i.e. there exists an  $\alpha \geq 0$  such that*

$$|B(u, v)| \leq \alpha \|u\| \|v\| \quad \text{for all } u, v \in H$$

*and*

(ii) *coercive, i.e. there exists a  $\beta > 0$  such that*

$$B(u, u) \geq \beta \|u\|^2 \quad \text{for all } u \in H.$$

*Then for any  $f \in H^*$  there exists a unique  $u_f \in H$  such that*

$$B(u_f, v) = f(v) \quad \text{for all } v \in H. \quad (1.20)$$

*Furthermore*

$$\|u_f\| \leq \beta^{-1} \|f\|_*; \quad (1.21)$$

*in particular  $u_f$  depends continuously on  $f$ , i.e.*

$$\|u_f - u_g\| \leq \beta^{-1} \|f - g\|_*.$$

*Proof* Once we have a solution it is clearly unique: if

$$B(u, v) = B(\bar{u}, v) = f(v)$$

for every  $v \in H$  then  $B(u - \bar{u}, v) = 0$  for every  $v \in H$  (since  $B$  is bilinear) and in particular for  $v = u - \bar{u}$ , whence

$$\beta \|u - \bar{u}\|^2 \leq B(u - \bar{u}, u - \bar{u}) = 0,$$

i.e.  $u = \bar{u}$ . The bound in (1.21) follows as before (set  $v = u_f$  in (1.20)) and the continuity result follows by considering

$$B(u_f, v) - B(u_g, v) = B(u_f - u_g, v) = (f - g, v)$$

and setting  $v = u_f - u_g$ . So only existence requires any work.

Fix  $u \in H$ , and consider the map  $v \mapsto B(u, v)$ . We claim that this defines a bounded linear functional on  $H$ : it is clearly linear, since

$$B(u, \alpha v_1 + \beta v_2) = \alpha B(u, v_1) + \beta B(u, v_2)$$

be the linearity of  $B$ , and it is bounded since

$$|B(u, v)| \leq \left[ \alpha \|u\| \right] \|v\|.$$

It follows from the Riesz Representation Theorem that there exists a  $w \in H$  such that

$$(w, v) = B(u, v) \quad \text{for all } v \in H.$$

We *define*  $Au = w$ , i.e. by definition

$$(Au, v) = B(u, v),$$

and claim that this definition yields a bounded linear operator from  $H$  into itself.

Indeed, for every  $v \in H$ ,

$$\begin{aligned} (A(\alpha u_1 + \beta u_2), v) &= B(\alpha u_1 + \beta u_2, v) \\ &= \alpha B(u_1, v) + \beta B(u_2, v) \\ &= \alpha (Au_1, v) + \beta (Au_2, v) \\ &= (\alpha Au_1 + \beta Au_2, v). \end{aligned}$$

Since this holds for every  $v \in H$ , it follows<sup>1</sup> that

$$A(\alpha u_1 + \beta u_2) = \alpha Au_1 + \beta Au_2,$$

i.e.  $A$  is linear. To show that  $A$  is bounded, note that

$$\|Au\|^2 = (Au, Au) = B(u, Au) \leq \alpha \|u\| \|Au\|,$$

so that  $\|Au\| \leq \alpha \|u\|$  and  $A$  is bounded.

Using the Riesz Representation Theorem in a standard way, we know that there exists a  $\varphi \in H$  such that

$$(\varphi, v) = f(v) \quad \text{for all } v \in H.$$

We can therefore rewrite our equation as

$$(Au, v) = (\varphi, v) \quad \text{for all } v \in H.$$

This implies that  $u$  satisfies (1.20) iff  $Au = \varphi$ ; we have regained a linear equation from the formulation in terms of a bilinear form. We now have to show that this equation has a solution.

<sup>1</sup> If  $(u, v) = (\bar{u}, v)$  for every  $v \in H$ , then  $(u - \bar{u}, v) = 0$  for every  $v \in H$ , in particular for  $v = u - \bar{u}$ , whence  $u = \bar{u}$ .

*Method 1: coercivity implies that  $A$  is invertible*

Now, not only is  $A$  ‘bounded above’; the coercivity implies that it is also bounded below:

$$\beta\|u\|^2 \leq B(u, u) = (Au, u) \leq \|Au\|\|u\| \quad \Rightarrow \quad \beta\|u\| \leq \|Au\|.$$

As a consequence  $A$  is one-to-one (if  $Au = Av$  then  $u = v$  since  $\|u - v\| \leq \beta^{-1}\|Au - Av\|$ ) and onto. To show that  $A$  is onto, first note that

$$R(A) = \{Au : u \in H\}$$

is a closed linear subspace of  $H$ . It is clearly a linear subspace, and it is closed since if  $v_n \in R(A)$  (so that  $v_n = Au_n$ ,  $u_n \in H$ ) and  $v_n \rightarrow v$ , then

$$\|u_n - u_m\| \leq \beta^{-1}\|Au_n - Au_m\| = \beta^{-1}\|v_n - v_m\|,$$

so that  $\{u_n\}$  is Cauchy. Since  $H$  is complete,  $u_n \rightarrow u \in H$ , and since  $A$  is bounded it is continuous: it follows that  $Au = v$ , i.e.  $v \in R(A)$ , so  $R(A)$  is closed.

So now suppose that  $R(A) \neq H$ . It follows that there exists a non-zero  $x \in H$  with  $x \in (R(A))^\perp$ . Thus

$$\beta\|x\|^2 \leq B(x, x) = (Ax, x) = 0,$$

a contradiction. Since  $A$  is one-to-one and onto, it must have an inverse, i.e. we can find a solution of  $Au = \varphi$  for any  $\varphi \in H$ .

*Method 2: contraction mapping argument*

Clearly, for any  $\varrho > 0$ ,

$$Au = \varphi \quad \Leftrightarrow \quad u = u - \varrho(Au - \varphi).$$

We use the contraction mapping theorem<sup>1</sup>, applied to the map

$$Tu = u - \varrho(Au - \varphi).$$

We have

$$\begin{aligned} \|Tu - Tv\|^2 &= \|(u - v) - \varrho A(u - v)\|^2 \\ &= \|u - v\|^2 - 2\varrho(A(u - v), u - v) + \varrho^2\|A(u - v)\|^2 \\ &= \|u - v\|^2 - 2\varrho B(u - v, u - v) + \varrho^2\|A(u - v)\|^2 \\ &\leq \|u - v\|^2 - 2\varrho\beta\|u - v\|^2 + \varrho^2\alpha^2\|u - v\|^2 \\ &= (1 - 2\varrho\beta + \varrho^2\alpha^2)\|u - v\|^2, \end{aligned}$$

<sup>1</sup> Let  $(X, d)$  be a complete metric space, and  $T : X \rightarrow X$  a map such that  $d(Tx, Ty) \leq \theta d(x, y)$  for some  $\theta < 1$ . Then  $T$  has a unique fixed point, i.e. there exists a unique  $x \in X$  such that  $Tx = x$ ; furthermore for any initial  $x_0$ , the iterates  $T^k x_0$  converge to the fixed point  $x$ . [To prove this, show that  $\{T^k x\}$  forms a Cauchy sequence, and use the completeness of  $X$ ; the limit is the fixed point.]

where we have used the coercivity of  $a$  and the fact that  $A$  is bounded. It follows that if we choose  $\varrho$  sufficiently small,  $T$  is a contraction. It therefore has a unique fixed point, which provides our solution.  $\square$

## 2

---

### Function spaces

We now need to introduce various function spaces so that we can perform our analysis rigorously.

#### 2.1 Euclidean spaces

$\mathbb{R}^n$  consists of all  $n$ -tuples  $\underline{x} = (x_1, \dots, x_n)$ . We use the standard Euclidean norm

$$|\underline{x}| = \left( \sum_{j=1}^n |x_j|^2 \right)^{1/2}$$

and inner product

$$\underline{x} \cdot \underline{y} = \sum_{j=1}^n x_j y_j.$$

$\mathbb{R}^n$  is complete.

#### 2.2 The sequence spaces $\ell^p$

The sequence spaces  $\ell^p$ ,  $1 \leq p < \infty$  consist of all sequences  $\underline{x} = (x_1, x_2, \dots)$  for which the  $\ell^p$  norm

$$\|\underline{x}\|_{\ell^p} := \left( \sum_{j=1}^{\infty} |x_j|^p \right)^{1/p}$$

is finite. In  $\ell^2$  the norm can be derived from the inner product

$$(\underline{x}, \underline{y}) = \sum_{j=1}^{\infty} x_j y_j.$$

The space  $\ell^\infty$  consists of all bounded sequences, with norm

$$\|\underline{x}\|_{\ell^\infty} := \sup_j |x_j|; \quad (2.1)$$

this space is slightly ‘odd’ (as, in some ways, is  $\ell^1$ ). Some work is required to show that these are norms.

We say that two indices  $1 \leq p, q \leq \infty$  are *conjugate* if

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (2.2)$$

The following simple inequality is fundamental.

**Lemma 2.1 (Young’s inequality)** *Let  $a, b > 0$  and let  $(p, q)$  be conjugate indices with  $1 < p, q < \infty$ . Then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \quad (2.3)$$

*Proof* The function  $e^x$  is convex<sup>1</sup>, i.e.  $e^{\lambda x + (1-\lambda)y} \leq \lambda e^x + (1-\lambda)e^y$ , and so

$$\begin{aligned} ab &= \exp(\log a + \log b) = \exp\left(\frac{1}{p} \log a^p + \frac{1}{q} \log b^q\right) \\ &\leq \frac{1}{p} e^{\log(a^p)} + \frac{1}{q} e^{\log(b^q)} = \frac{a^p}{p} + \frac{b^q}{q}. \end{aligned}$$

□

**Lemma 2.2 (Hölder’s inequality in  $\ell^p$  spaces)** *Let  $\underline{x} \in \ell^p$  and  $\underline{y} \in \ell^q$  with  $p, q$  conjugate,  $1 \leq p, q \leq \infty$ . Then if  $\underline{z} = (x_1 y_1, x_2 y_2, \dots)$ ,  $\underline{z} \in \ell^1$  with*

$$\|\underline{z}\|_{\ell^1} = \sum_{j=1}^{\infty} |x_j y_j| \leq \|\underline{x}\|_{\ell^p} \|\underline{y}\|_{\ell^q}. \quad (2.4)$$

<sup>1</sup> A twice differentiable function on an interval  $(a, b)$  is convex on  $(a, b)$  iff its second derivative is non-negative, see examples.

*Proof* For  $1 < p < \infty$ , consider

$$\sum_{j=1}^n \frac{|x_j|}{\|\underline{x}\|_{\ell^p}} \frac{|y_j|}{\|\underline{y}\|_{\ell^q}} \leq \sum_{j=1}^n \frac{1}{p} \frac{|x_j|^p}{\|\underline{x}\|_{\ell^p}^p} + \frac{1}{q} \frac{|y_j|^q}{\|\underline{y}\|_{\ell^q}^q} \leq 1.$$

So for each  $n \in \mathbb{N}$

$$\sum_{j=1}^n |x_j y_j| \leq \|\underline{x}\|_{\ell^p} \|\underline{y}\|_{\ell^q}$$

and (2.4) follows. For  $p = 1$ ,  $q = \infty$ ,

$$\sum_{j=1}^n |x_j y_j| \leq \max_{j=1}^n |y_j| \left( \sum_{j=1}^n |x_j| \right) \leq \|\underline{x}\|_{\ell^1} \|\underline{y}\|_{\ell^\infty}.$$

□

**Lemma 2.3 (Minkowski's inequality in  $\ell^p$  spaces)** *If  $\underline{x}, \underline{y} \in \ell^p$  then  $\underline{x} + \underline{y} \in \ell^p$  and*

$$\|\underline{x} + \underline{y}\|_{\ell^p} \leq \|\underline{x}\|_{\ell^p} + \|\underline{y}\|_{\ell^p}.$$

*Proof* The cases  $p = 1, \infty$  are easy. For  $1 \leq p < \infty$  let  $q$  be the conjugate exponent to  $p$ , and note that  $(p-1)q = p$ . Since  $|x_j + y_j|^p \leq 2^p(|x_j|^p + |y_j|^p)$  clearly  $\underline{x} + \underline{y} \in \ell^p$ ; now use Hölder's inequality to write

$$\begin{aligned} \sum_{j=1}^n |x_j + y_j|^p &\leq \sum_{j=1}^n |x_j + y_j|^{p-1} |x_j| + \sum_{j=1}^n |x_j + y_j|^{p-1} |y_j| \\ &\leq \left\{ \sum_{j=1}^n |x_j + y_j|^{(p-1)q} \right\}^{1/q} \left\{ \left( \sum_{j=1}^n |x_j|^p \right)^{1/p} + \left( \sum_{j=1}^n |y_j|^p \right)^{1/p} \right\} \\ &\leq \left( \sum_{j=1}^n |x_j + y_j|^p \right)^{1/q} (\|\underline{x}\|_{\ell^p} + \|\underline{y}\|_{\ell^p}), \end{aligned}$$

and so

$$\left( \sum_{j=1}^n |x_j + y_j|^p \right)^{1-(1/q)} \leq \|\underline{x}\|_{\ell^p} + \|\underline{y}\|_{\ell^p},$$

from whence the triangle inequality in  $\ell^p$  follows. □

**Proposition 2.4** *For each  $1 \leq p \leq \infty$ , the sequence space  $\ell^p$  (equipped with its standard norm) is complete.*

For a proof see Problems 2.

Thus every  $\ell^p$  is a Banach space, and  $\ell^2$  is a Hilbert space. Every (infinite-dimensional) separable Hilbert space is isometrically isomorphic to  $\ell^2$ .

Since the norm on  $\ell^2$  is the natural generalisation of the norm on  $\mathbb{R}^n$ , and since it is complete, it is tempting to think that  $\ell^2$  will behave just like  $\mathbb{R}^n$ . However, it does not have the ‘Bolzano-Weierstrass property’ (bounded sequences have a convergent subsequence) as we can see easily by considering the sequence  $\{\underline{e}_j\}_{j=1}^{\infty}$ , where  $\underline{e}_j$  consists entirely of zeros apart from a 1 in the  $j$ th position. Then clearly  $\|\underline{e}_j\|_{\ell^2} = 1$  for all  $j$ ; but if  $i \neq j$  then

$$\|\underline{e}_i - \underline{e}_j\|_{\ell^2}^2 = 2,$$

i.e. any two elements of the sequence are always  $\sqrt{2}$  away from each other. It follows that no subsequence of the  $\{\underline{e}_j\}$  can form a Cauchy sequence, and so there cannot be a convergent subsequence.

This is really the first time we have seen a significant difference between  $\mathbb{R}^n$  and the abstract normed vector spaces that we have been considering. The failure of the Bolzano-Weierstrass property is in fact a defining characteristic of infinite-dimensional spaces.

### 2.3 Density and separability

A set  $A$  is *dense* in a normed space  $(X, \|\cdot\|)$  if every  $x \in X$  can be approximated arbitrarily closely by an element of  $A$ : given  $x \in X$  and  $\epsilon > 0$  there exists an  $a \in A$  such that

$$\|x - a\| < \epsilon.$$

When  $X$  is complete, this is equivalent to requiring that  $X = \overline{A}$ , the closure of  $A$  (wrt the norm  $\|\cdot\|$ ).

If a space has a countable dense subset, it is called separable.

$\mathbb{R}^n$  and  $\ell^p$ ,  $1 \leq p < \infty$  are separable; however,  $\ell^\infty$  is *not* separable.

**Proposition 2.5**  *$\ell^p$ ,  $1 \leq p < \infty$ , is separable.*

*Proof* Let  $\mathbb{Q}^\sigma$  be the collection of sequence that have only a finite number of non-zero terms, each of which is an element of  $\mathbb{Q}$ . Then  $\mathbb{Q}^\sigma$  is countable (it is a countable union of countable sets, i.e. those with  $\leq n$  non-zero terms), and dense in  $\ell^p$ : given  $\underline{x} \in \ell^p$  and  $\epsilon > 0$ , choose  $N$  such that

$$\sum_{j=N+1}^{\infty} |x_j|^p < \epsilon^p/2,$$

and then choose  $\underline{y} \in \mathbb{Q}^\sigma$  such that  $y_j = 0$  for  $j \geq N+1$  and  $|y_j - x_j| < \epsilon^p/2N$  for  $j < N$ . It follows that  $\|\underline{x} - \underline{y}\|_{\ell^p} < \epsilon$ .  $\square$

**Proposition 2.6**  $\ell^\infty$  is not separable.

*Proof* Suppose that  $A$  is dense in  $\ell^\infty$ . Note that the uncountable set  $L = \{0, 1\}^\infty$  is a subset of  $\ell^\infty$ . Choose  $\epsilon < 1/2$ ; then we claim that for every element of  $L$  there exists a distinct element of  $A$ ; indeed, if  $\underline{x}, \underline{y} \in L$  with  $\underline{x} \neq \underline{y}$  then there exists an index  $j$  such that  $x_j \neq y_j$ ; wlog  $x_j = 1$  and  $y_j = 0$ . It follows that if  $\underline{a}, \underline{b} \in A$  with

$$\|\underline{a} - \underline{x}\|_{\ell^\infty} < \epsilon \quad \text{and} \quad \|\underline{b} - \underline{y}\|_{\ell^\infty} < \epsilon$$

then  $a_j > 1/2$  and  $b_j < 1/2$ , i.e.  $a_j \neq b_j$  and hence  $\underline{a} \neq \underline{b}$ . It follows that  $A$  is uncountable, and hence  $\ell^\infty$  cannot be separable.  $\square$

The subspace  $c_0$  of  $\ell^\infty$ , consists of all sequences that converge to zero,

$$c_0 = \{\underline{x} : x_j \rightarrow 0\}$$

equipped with the  $\ell^\infty$  norm is still complete, and is separable.

## 2.4 Spaces of continuous and differentiable functions

Let  $\Omega$  be an open subset of  $\mathbb{R}^n$  (which could be  $\mathbb{R}^n$  itself). We will be interested in a variety of spaces of continuous functions (and functions with continuous derivatives). Note that not all of these spaces are complete.

### 2.4.1 Spaces of continuous functions

The space  $C^0(\Omega)$  consists of all (real-valued) continuous functions on  $\Omega$ .

The space  $C^0(\bar{\Omega})$  consists of all real-valued continuous functions on  $\bar{\Omega}$ , the closure of  $\Omega$ .

**Theorem 2.7** *The space  $C^0(\bar{\Omega})$  is complete when equipped with the supremum norm,*

$$\|f\|_{\infty} = \sup_{x \in \bar{\Omega}} |f(x)| = \sup_{x \in \Omega} |f(x)|.$$

*Proof* Let  $\{f_n\}$  be a Cauchy sequence wrt the sup norm. Then given  $\epsilon > 0$ , there exists an  $N$  such that

$$\sup_{x \in \bar{\Omega}} |f_n(x) - f_m(x)| < \epsilon \quad \text{for all } n, m \geq N.$$

It follows that for each  $x \in \bar{\Omega}$ ,

$$|f_n(x) - f_m(x)| < \epsilon \quad \text{for all } n, m \geq N, \quad (2.5)$$

i.e.  $\{f_n(x)\}$  is a Cauchy sequence (but note that  $N$  does not depend on  $x$ ). Define a function  $f : \bar{\Omega} \rightarrow \mathbb{R}$  by

$$f(x) = \lim_{n \rightarrow \infty} f_n(x).$$

Then  $f$  is in fact the uniform limit on the  $f_k$ ; letting  $m \rightarrow \infty$  in (2.5) shows that for each fixed  $x \in \bar{\Omega}$ ,

$$|f_n(x) - f(x)| < \epsilon \quad \text{for all } n \geq N,$$

where again  $N$  does not depend on  $x$ , whence

$$\sup_{x \in \bar{\Omega}} |f_n(x) - f(x)| < \epsilon \quad \text{for all } n \geq N.$$

It is then standard that  $f$  (the uniform limit of continuous functions) is continuous itself.  $\square$

The space  $C^0(\Omega)$  is *not* complete with the sup norm.

Using the Weierstrass Approximation Theorem (polynomials are dense in the space of continuous functions on a compact set, see Problems sheet), it follows immediately that:

**Theorem 2.8** *The space  $C^0(\bar{\Omega})$  is separable.*

### 2.4.2 Multi-index notation and spaces of differentiable functions

A  $n$ -dimensional multi-index is an  $n$ -tuple  $\alpha = (\alpha_1, \dots, \alpha_n)$  of integers. We write

$$|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n.$$

If  $x \in \mathbb{R}^n$  is a vector; or  $\partial = (\partial_1, \partial_2, \dots, \partial_n)$  is a vector of operators (we abbreviate  $\partial/\partial x_j$  to  $\partial_j$ ) then

$$x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}, \quad \partial^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

For example, in  $\mathbb{R}^2$ ,

$$\sum_{|\alpha|=2} \partial^\alpha f = \frac{\partial^2 f}{\partial x_1^2} + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\partial^2 f}{\partial x_2^2};$$

and Leibniz's rule becomes

$$\partial^\alpha (fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} f \partial^\beta g,$$

where

$$\binom{\alpha}{\beta} = \frac{\alpha!}{(\alpha-\beta)! \beta!},$$

$\alpha! = \alpha_1! \dots \alpha_n!$ , and  $\beta \leq \alpha$  if  $\beta_i \leq \alpha_i$  for all  $i = 1, \dots, n$ .

The space  $C^k(\Omega)$  consists of all real-valued functions

$$\{f : \Omega \rightarrow \mathbb{R} \mid \partial^\alpha f \in C^0(\Omega), |\alpha| \leq k\},$$

i.e. all derivatives up to order  $k$  exist and are continuous on  $\Omega$ .

With a similar definition for  $C^k(\bar{\Omega})$ , it is relatively straightforward to show that  $C^k(\bar{\Omega})$  is complete wrt the  $C^k$  norm

$$\|f\|_{C^k} = \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_\infty.$$

### 2.4.3 'Test functions'

Extremely useful are functions that are infinitely differentiable and have compact support. The support of a function  $f : \Omega \rightarrow \mathbb{R}$  is the smallest

closed set containing

$$\{x \in \Omega : f(x) \neq 0\}.$$

We write  $C^\infty(\Omega) = \bigcap_{k \geq 0} C^k(\Omega)$  [and similarly for  $\bar{\Omega}$ ]. Then space  $C_c^\infty(\Omega)$  of ‘test functions in  $\Omega$  (beware, this is not a terminology that is used in any way consistently) consists of all  $C^\infty$  functions with compact support in  $\Omega$  – since  $\Omega$  is open, the support of  $f$  will be bounded away from  $\partial\Omega$  (and this is often extremely helpful).

One could make similar definitions with less differentiability, e.g.  $C_c^0(\Omega)$ , continuous functions with compact support in  $\Omega$ . This space is not complete with the sup norm.

#### 2.4.4 Mollification

We now introduce the important idea of *mollification*. Essentially this is a way of obtaining a very smooth function from a less smooth function by averaging. We choose a non-negative function  $\rho \in C_c^\infty(\mathbb{R}^n)$  whose support is contained in  $B(0, 1) = \{x : |x| \leq 1\}$  and which integrates to one, for example

$$\rho(x) = \left\{ c \exp\left(\frac{1}{|x|^2-1}\right), \quad |x| \leq 1, 0, \quad |x| \geq 1, \right\},$$

where  $c$  is chosen such that

$$\int_{\mathbb{R}^n} \rho(x) \, dx = 1.$$

We write  $\rho_h(x) = h^{-n} \rho(x/h)$  (the support of  $\rho_h$  is contained in a ball about the origin of radius  $h$ , and it still integrates to 1), and define the mollification of  $f$ ,  $f_h$ , as the convolution of  $\rho_h$  with  $f$ ,

$$f_h(x) = (\rho_h \star f)(x) = h^{-n} \int_{\mathbb{R}^n} \rho\left(\frac{x-y}{h}\right) f(y) \, dy.$$

[If we start with a function  $u \in C^0(\Omega)$ , say, then we simply extend  $u$  by zero outside  $\Omega$ ; equivalently, we restrict the range of integration to  $\Omega$ .]

**Proposition 2.9 (Mollification of functions in  $C_c^0(\Omega)$ .)** *Let  $f \in C_c^0(\Omega)$ . Then  $f_h \in C_c^\infty(\Omega)$  if  $h < \text{dist}(\text{supp } u, \partial\Omega)$ , and  $f_h \rightarrow f$  uniformly in  $\Omega$  as  $h \rightarrow 0$ .*

*Proof* If  $h < \text{dist}(\text{supp } f, \partial\Omega)$  then (i) we can differentiate under the integral sign with respect to  $x$  to show that  $f_h$  is in  $C^\infty$ , and (ii) the support of  $f_h$  must be contained within an  $h$ -neighbourhood of the support of  $f$ , and hence within a compact subset of  $\Omega$ .

To show convergence note that

$$f_h(x) - f(x) = h^{-n} \int \rho\left(\frac{x-y}{h}\right) [f(y) - f(x)] dy,$$

since  $\int \rho_h = 1$ . Then clearly

$$|f_h(x) - f(x)| \leq \sup_{|y-x| \leq h} |f(y) - f(x)|.$$

Since the support of  $f$  is compact,  $f$  is uniformly continuous on  $\Omega$ ; given an  $\epsilon > 0$  there exists a  $\delta > 0$  such that  $|x-y| < \delta$  implies that  $|f(x) - f(y)| < \epsilon$ , and hence for  $h < \delta$ ,  $|f_h(x) - f(x)| < \epsilon$ , uniformly over  $\Omega$ .  $\square$

## 2.5 The Lebesgue integral on $\mathbb{R}$

Here is a very cursory outline of the construction of the Lebesgue integral on  $\mathbb{R}$ . For more details see Priestley's book *Introduction to Integration* or my *Functional Analysis I* notes.

A subset  $U$  of  $\mathbb{R}$  has *zero measure* if for every  $\epsilon > 0$  there exists a collection of intervals  $(a_j, b_j)$  such that

$$U \subset \bigcup_{j=1}^{\infty} (a_j, b_j)$$

and

$$\sum_{j=1}^{\infty} |b_j - a_j| < \epsilon.$$

The class  $L^{\text{step}}(\mathbb{R})$  of *step functions* consists of all those functions  $s(x)$  that are non-negative and piecewise constant on a finite number of intervals, i.e.

$$s(x) = \sum_{j=1}^n c_j \chi[I_j](x),$$

where  $c_j \geq 0$ , each  $I_j$  is an interval, and  $\chi[A]$  denotes the characteristic

function of  $A$ ,

$$\chi[A](x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

We define the integral of  $s(x)$  by

$$\int s = \sum_{j=1}^n c_j |I_j|,$$

where  $|I_j| = (b_j - a_j)$ . It is tedious but elementary to check that this gives a well-defined integral on  $L^{\text{step}}(\mathbb{R})$ .

Now, if  $\{s_n\}$  is a monotonically increasing sequence of functions in  $L^{\text{step}}(\mathbb{R})$ , i.e.  $s_{n+1}(x) \geq s_n(x)$  for each  $x \in \mathbb{R}$  then the sequence

$$\int s_n$$

is also monotonically increasing. So provided that these integrals are bounded above,

$$\lim_{n \rightarrow \infty} \int s_n$$

must exist. A key point – whose proof is fairly technical – is that under these conditions, the functions  $s_n$  converge *almost everywhere*, i.e. except on a set of measure zero, to a function  $f$ .

Denote the set of all functions that can be produced in this way (the almost everywhere limit of an increasing sequence of step functions whose integrals are uniformly bounded above) by  $L^{\text{inc}}(\mathbb{R})$ ; for every  $f \in L^{\text{inc}}(\mathbb{R})$ , with  $f(x) = \lim_{n \rightarrow \infty} s_n(x)$  almost everywhere, we define

$$\int f = \lim_{n \rightarrow \infty} \int s_n.$$

Again, one has to show that this definition is unambiguous (and this is not straightforward).

Finally, we let  $L^1(\mathbb{R})$  be the collection of all functions  $f = g - h$ , where  $g, h \in L^{\text{inc}}(\mathbb{R})$ , and for such an  $f$  define

$$\int f = \int g - \int h.$$

We have to check that this is well-defined (this time it's easy), and can then show that this integral – the Lebesgue integral – has the following desirable properties.

(L) (Linearity) If  $f_1, f_2 \in L^1(\mathbb{R})$  and  $\lambda \in \mathbb{R}$  then  $f_1 + \lambda f_2 \in L^1(\mathbb{R})$  and

$$\int f_1 + \lambda f_2 = \int f_1 + \lambda \int f_2.$$

(P) (Positivity) If  $f \in L^1(\mathbb{R})$  and  $f \geq 0$  almost everywhere then  $\int f \geq 0$ .

(M) (Modulus property) If  $f \in L^1(\mathbb{R})$  then<sup>1</sup>  $|f| \in L^1(\mathbb{R})$  and

$$\left| \int f \right| \leq \int |f|.$$

(T) (Translation invariance) If  $f \in L^1(\mathbb{R})$  and  $f_d(x) = f(x + d)$  then  $f_d \in L^1(\mathbb{R})$  and  $\int_{\mathbb{R}} f = \int_{\mathbb{R}} f_d$ .

**Theorem 2.10 (Monotone Convergence Theorem)** *If  $f_n \in L^1(\mathbb{R})$  with*

$$0 \leq f_1(x) \leq f_2(x) \leq \dots \quad \text{for a.e. } x \in \mathbb{R}$$

and

$$\int f_n \leq M$$

for some  $M > 0$  then there exists an  $f \in L^1(\mathbb{R})$  such that  $f_n(x)$  converges to  $f(x)$  almost everywhere and

$$\int f = \lim_{n \rightarrow \infty} \int f_n.$$

Suppose that  $f \geq 0$  and  $\int f = 0$ . Consider  $f_n(x) = nf(x)$ ; then  $f_{n+1} \geq f_n$ , and  $\int f_n = n \int f = 0$ , so by the MCT  $f_n(x) \rightarrow g(x)$  almost everywhere, with  $g(x) \in L^1$ . Since  $g(x) = +\infty$  where  $f \neq 0$ , and  $g$  is defined almost everywhere (since it is in  $L^1$ ), it follows that  $f = 0$  almost everywhere. This shows that if  $\int |f| = 0$  then one can only deduce that  $f = 0$  almost everywhere.

**Theorem 2.11 ((Lebesgue's) Dominated Convergence theorem)** *If  $f_n \in L^1(\mathbb{R})$  converge pointwise almost everywhere to  $f$ , and there is a function  $g \in L^1(\mathbb{R})$  such that  $|f_n(x)| \leq g(x)$  for every  $n$  and almost every  $x \in \mathbb{R}$ , then  $f \in L^1(\mathbb{R})$  and*

$$\int f = \lim_{n \rightarrow \infty} \int f_n.$$

<sup>1</sup> This does imply that there are some functions whose 'integral' exists in some sense, e.g.  $\int_1^\infty \frac{1}{x} \sin x \, dx$ , which are not in  $L^1(\mathbb{R})$ .

We say that  $f \in L^1(a, b)$  if  $f|_{(a,b)} = f\chi_{(a,b)} \in L^1(\mathbb{R})$ .

Note that any integrable function  $f \in L^1(\mathbb{R})$  is given as  $f = g - h$  where  $g, h \in L^{\text{inc}}(\mathbb{R})$ . We have  $g = \lim_{n \rightarrow \infty} g_n$  and  $h = \lim_{n \rightarrow \infty} h_n$ , where  $\{g_n\}$  and  $\{h_n\}$  are monotonic sequences of functions in  $L^{\text{step}}$ , converging almost everywhere. This implies, trivially, that  $f$  is the limit almost everywhere of the (not necessarily monotonic) sequence of step functions  $f_n - g_n$ . Such a function (the almost everywhere limit of a sequence of step functions) is called *measurable*.

## 2.6 The Lebesgue spaces $L^p$

The Lebesgue space  $L^p(\Omega)$  consists of all those functions that are measurable and whose  $p$ th power is Lebesgue integrable;

$$L^p(\Omega) = \left\{ f : f \text{ is measurable and } \int_{\Omega} |f|^p < \infty \right\},$$

and is a Banach space when equipped with the  $L^p$  norm

$$\|f\|_{L^p} = \left( \int_{\Omega} |f|^p \right)^{1/p}. \quad (2.6)$$

We now show (i) that this is a norm and (ii) that  $L^p$  is complete wrt this norm.

The space  $L^\infty(\Omega)$  consists of all measurable functions on  $\Omega$  such that

$$\|f\|_{L^\infty} := \text{ess sup}_{x \in \Omega} |f(x)| < \infty;$$

where the essential supremum is

$$\inf\{M : |f(x)| \leq M \text{ for all } x \in \Omega \setminus E, \text{ where } E \text{ has measure zero}\}.$$

**Lemma 2.12 (Hölder's inequality)** *Let  $1 \leq p, q \leq \infty$  be conjugate indices. Suppose that  $f \in L^p$  and  $g \in L^q$ . Then  $fg \in L^1$  and*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}. \quad (2.7)$$

*Proof* For  $1 < p, q < \infty$ ,

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

For  $p = 1$ ,  $q = \infty$  the inequality is clear:

$$\int |f(x)||g(x)| \leq \|g\|_{L^\infty} \int |f(x)| = \|f\|_{L^1} \|g\|_{L^\infty}.$$

□

One can generalise Hölder's inequality to treat the product of three or more functions. The three-term version is sometimes useful: if  $f \in L^p$ ,  $g \in L^q$ , and  $h \in L^r$  with  $p^{-1} + q^{-1} + r^{-1} = 1$ , then  $fgh \in L^1$  with

$$\int |fgh| \leq \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}.$$

For a proof see the examples sheet.

The  $L^p$  interpolation inequality is also useful: take  $p < r < q$  with

$$\frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q};$$

then if  $f \in L^p$  and  $f \in L^q$ ,

$$\|f\|_{L^r} \leq \|f\|_{L^p}^\alpha \|f\|_{L^q}^{1-\alpha}.$$

Again, the proof is a simple application of Hölder's inequality (see examples).

**Lemma 2.13 (Minkowski's inequality)** For  $1 \leq p \leq \infty$ ,  $f \in L^p$  and  $g \in L^p$  then  $f + g \in L^p$  and

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

*Proof* First note that  $|f(x) + g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p)$ , since  $f + g \in L^p$ . Also, if  $(p, q)$  are conjugate then  $(p-1)q = p$ ; so we can apply Hölder's inequality to obtain

$$\begin{aligned} \|f + g\|_{L^p}^p &= \int |f + g|^p \leq \int |f| |f + g|^{p-1} + \int |g| |f + g|^{p-1} \\ &\leq \|f\|_{L^p} \|f + g\|_{L^p}^{p/q} + \|g\|_{L^p} \|f + g\|_{L^p}^{p/q} \\ &\leq (\|f\|_{L^p} + \|g\|_{L^p}) \|f + g\|_{L^p}^{p/q}; \end{aligned}$$

dividing both sides by  $\|f + g\|_{L^p}^{p/q}$  yields the result, since

$$p - \frac{p}{q} = p \left(1 - \frac{1}{q}\right) = 1.$$

□

We now show that  $L^p$  is complete.

**Theorem 2.14**  $L^p(\mathbb{R})$  is complete.

*Proof* Suppose that  $\{f_n\}$  is a Cauchy sequence in  $L^p$ . Choose  $n_1 < n_2 < \dots$  such that

$$\|f_n - f_m\|_{L^p} < 2^{-k} \quad \text{for } n, m \geq n_k;$$

then the subsequence  $f_{n_j}$  satisfies

$$\|f_{n_{j+1}} - f_{n_j}\|_{L^p} < 2^{-j}.$$

It follows that

$$\left\| \sum_{j=1}^k |f_{n_{j+1}} - f_{n_j}| \right\|_{L^p} \leq \sum_{j=1}^k \|f_{n_{j+1}} - f_{n_j}\|_{L^p} \leq \sum_{j=1}^k 2^{-j} \leq 1.$$

If we define

$$v_k(x) = \sum_{j=1}^k |f_{n_{j+1}}(x) - f_{n_j}(x)|,$$

then it follows from the Monotone Convergence Theorem that  $v_k \rightarrow v$  almost everywhere for some  $v \in L^p$ ; in particular the series

$$\sum_{j=1}^k f_{n_{j+1}}(x) - f_{n_j}(x)$$

converges absolutely almost everywhere. Since the partial sums of this series are just  $f_{n_{k+1}}(x) - f_{n_1}(x)$ , it follows that  $f_{n_j}(x)$  converges almost everywhere; at these points we set

$$f(x) = \lim_{j \rightarrow \infty} f_{n_j}(x)$$

and define  $f$  however we like elsewhere. All that remains is to show that  $f \in L^p(\mathbb{R})$  and that  $\|f_n - f\|_{L^p} \rightarrow 0$ .

Now consider

$$\begin{aligned}
 |f - f_{n_{k+1}}|^p &= \left| (f - f_{n_1}) - \sum_{j=1}^k (f_{n_{j+1}} - f_{n_j}) \right|^p \\
 &= \left| \sum_{j=k+1}^{\infty} (f_{n_{j+1}} - f_{n_j}) \right|^p \\
 &\leq \left( \sum_{j=k+1}^{\infty} |f_{n_{j+1}} - f_{n_j}| \right)^p \\
 &\leq |v|^p,
 \end{aligned}$$

where  $v \in L^p$ . It follows that  $f - f_{n_{k+1}} \in L^p$  and hence, since  $f_{n_{k+1}} \in L^p$ , that  $f \in L^p$ .

Finally, note that we can use the dominated convergence theorem to obtain

$$\begin{aligned}
 \|f - f_{n_{k+1}}\|_{L^p} &= \left\| f - f_{n_1} - \sum_{j=1}^k (f_{n_{j+1}} - f_{n_j}) \right\|_{L^p} \\
 &= \left\| \sum_{j=k+1}^{\infty} (f_{n_{j+1}} - f_{n_j}) \right\|_{L^p} \\
 &\leq \sum_{j=k+1}^{\infty} \|f_{n_{j+1}} - f_{n_j}\|_{L^p} \\
 &\leq 2^{-(k-1)},
 \end{aligned}$$

so that  $f \in L^p$  (since  $f = (f - f_{n_{k+1}}) + f_{n_{k+1}}$  and both the terms on the right-hand side are in  $L^p$ ) and  $f_{n_k} \rightarrow f$  in  $L^p$ . We have already seen (in the first problems sheet) that if a subsequence of a Cauchy sequence converges then the entire sequence converges, and the proof is complete.  $\square$

The following useful corollary is contained in the proof.

**Corollary 2.15** *Suppose that  $f_n \rightarrow f$  in  $L^p$ . Then there is a subsequence  $\{f_{n_j}\}$  such that  $f_{n_j} \rightarrow f$  almost everywhere.*

The following observation, which follows from the definition of  $L^1$ , is fundamental.

**Lemma 2.16** *The space  $C_c^0(\Omega)$  is dense in  $L^p(\Omega)$  for  $1 \leq p < \infty$ .*

*Proof (Sketch)* Given  $f \in L^p(\Omega)$  and  $\epsilon > 0$  find a compact subset  $K$  of  $\Omega$  such that

$$\|f - f\chi_K\|_{L^p} < \epsilon/2.$$

The problem reduces to approximating  $f\chi_K \in L^p(K)$ . Write  $f\chi_K = f_1 - f_2$  with  $f_1, f_2 \in L^{\text{inc}}(K)$ . The problem reduces to approximating  $f_j$  to within  $\epsilon/4$ .

If  $f \in L^{\text{inc}}$  then it is the limit of an increasing sequence  $s_n$  of step functions. Since  $s_n(x) \leq f(x)$  it follows that  $s_n \in L^p$ , and since  $(u(x) - s_n(x))^p \leq u(x)^p$  it follows that  $s_n \rightarrow u$  in  $L^p$  using the DCT.

Any step functions can be approximated arbitrarily closely in the  $L^p$  norm by continuous functions, which completes the proof.  $\square$

By mollification we immediately obtain:

**Corollary 2.17** *The space  $C_c^\infty(\Omega)$  is dense in  $L^p(\Omega)$  for  $1 \leq p < \infty$ .*

*Proof (Sketch)* The uniform norm on any compact set  $K$  dominates the  $L^p$  norm on  $K$ :

$$\int_K |f(x)|^p dx \leq |K| \|f\|_\infty^p.$$

So approximate  $f$  to within  $\epsilon/2$  by a function  $g \in C_c^0(\Omega)$ ; then mollify.  $\square$

Furthermore it follows from the separability of  $C^0(\bar{\Omega})$  that the ‘nice’  $L^p$  spaces are separable:

**Theorem 2.18**  *$L^p(\Omega)$  is separable for  $1 \leq p < \infty$ .*

*Proof (Sketch)* Write

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

where

$$\Omega_j = \{x \in \Omega : \text{dist}(x, \partial\Omega) > 1/j \text{ and } |x| < j\}.$$

Use the facts that (i)  $C^0(\bar{\Omega}_j)$  is separable for each  $j$ , (ii) the sup norm on  $\bar{\Omega}_j$

dominates the  $L^p(\bar{\Omega}_j)$  norm, and (iii) a countable union of countable sets is still countable.  $\square$

## 2.7 Locally integrable functions

We say that a function  $f$  is locally in  $L^p$  in  $\Omega$ , or  $f \in L^p_{\text{loc}}(\Omega)$ , if  $f \in L^p(K)$  for every compact subset  $K$  of  $\Omega$ . We say that  $f_n \rightarrow f$  'locally in  $L^p(\Omega)$ ' if  $f_n \rightarrow f$  in  $L^p(K)$  for every compact subset  $K$  of  $\Omega$  [usually  $\Omega = \mathbb{R}^n$  for such local spaces].

Recall that the mollification  $f_h$  of  $f$  is given by

$$f_h(x) = (\rho_h \star f)(x) = \int_{\mathbb{R}^n} \rho_h(x-y)f(y) \, dy = h^{-n} \int_{\mathbb{R}^n} \rho\left(\frac{x-y}{h}\right) f(y) \, dy. \quad (2.8)$$

**Lemma 2.19 (Mollification of functions in  $L^p_{\text{loc}}$ )** *If  $f \in L^p_{\text{loc}}(\Omega)$ ,  $1 \leq p < \infty$ , then  $f_h \in L^p_{\text{loc}}(\Omega)$  and  $f_h \rightarrow f$  in  $L^p_{\text{loc}}(\Omega)$ . Furthermore,  $f_h \in C^\infty(\Omega)$ , and if  $f$  has compact support in  $\Omega$  and  $h < \text{dist}(\text{supp } f, \partial\Omega)$  then  $f_h \in C_c^\infty(\Omega)$ .*

Why can this result not hold for  $p = \infty$ ?

*Proof* First we show that if  $K$  is a compact subset of  $\Omega$ , then there is another compact subset  $K'$  of  $\Omega$  such that

$$\|f_h\|_{L^p(K)} \leq \|f\|_{L^p(K')}.$$

Take  $h < \frac{1}{2} \text{dist}(K, \partial\Omega)$ , and then change variables in (2.8) with  $z = x - y$  to give

$$f_h(x) = \int_{|z| \leq h} \rho_h(z) f(x-z) \, dz.$$

Since  $\int \rho_h = 1$ , using Hölder's inequality

$$\begin{aligned}
|f_h(x)|^p &= \left( \int_{|z| \leq h} \rho_h(z) f(x-z) \, dz \right)^p \\
&= \left( \int_{|z| \leq h} \rho_h(z)^{1-(1/p)} \rho_h(z)^{1/p} f(x-z) \, dz \right)^p \\
&\leq \left( \int_{|z| \leq h} \rho_h(z)^{q(1-(1/p))} \, dz \right)^{p/q} \left( \int_{|z| \leq h} \rho_h(z) |f(x-z)|^p \, dz \right) \\
&= \int_{|z| \leq h} \rho_h(z) |f(x-z)|^p \, dz.
\end{aligned}$$

So

$$\begin{aligned}
\int_K |f_h(x)|^p \, dx &\leq \int_K \int_{|z| \leq h} \rho_h(z) |f(x-z)|^p \, dz \, dx \\
&= \int_{|z| \leq h} \rho_h(z) \int_K |f(x-z)|^p \, dx \, dz \\
&\leq \int_{|z| \leq h} \rho_h(z) \left( \int_{K'} |f(x)|^p \, dx \right) \\
&= \int_{K'} |f(x)|^p \, dx,
\end{aligned}$$

where

$$K' = \{y = x + z : x \in K, |z| \leq h\}.$$

To complete the proof we use the density of  $C_c^0(\Omega)$  in  $L^p(\Omega)$ , and mollification of functions in  $C_c^0(\Omega)$ : given  $\epsilon > 0$ , choose  $\phi \in C_c^0(\Omega)$  such that

$$\|f - \phi\|_{L^p(\Omega)} < \epsilon/3.$$

Choose  $h$  small enough that  $\|\phi - \phi_h\|_{L^p(K')} < \epsilon/3$  (the sup norm on  $K'$  dominates the  $L^p(K')$  norm); then

$$\begin{aligned}
\|f - f_h\|_{L^p(K)} &\leq \|f - \phi\|_{L^p(K)} + \|\phi - \phi_h\|_{L^p(K)} + \|\phi_h - f_h\|_{L^p(K)} \\
&\leq 2\epsilon/3 + \|\phi - f\|_{L^p(K')} \leq \epsilon.
\end{aligned}$$

□

## 2.8 Sobolev spaces

### 2.8.1 Weak derivatives

Suppose that  $f$  is differentiable on  $\Omega$ . Then for any test function  $\phi \in C_c^\infty(\Omega)$ , one can integrate by parts to obtain

$$\int_{\Omega} \frac{\partial f}{\partial x_j} \phi \, dx = - \int_{\Omega} f \frac{\partial \phi}{\partial x_j} \, dx, \quad (2.9)$$

using the fact that  $\phi$  has compact support in  $\Omega$  (so the boundary term vanishes). Repeating this process  $|\alpha|$  times yields

$$\int_{\Omega} (\partial^\alpha f) \phi \, dx = (-1)^{|\alpha|} \int_{\Omega} f \partial^\alpha \phi \, dx$$

for any multi-index  $\alpha$ .

Now, while the left-hand side of (2.9) makes sense only if  $\partial f/\partial x_j$  exists, the right-hand side makes sense for any  $f \in L^1$ , since we always have  $\partial_j \phi \in L^\infty$ . For a function  $f \in L^1_{\text{loc}}(\Omega)$ , we say that  $g$  is the weak derivative of  $f$  with respect to  $x_j$ , and write  $\partial_j f = g$ , if  $g \in L^1_{\text{loc}}(\Omega)$  and

$$\int_{\Omega} g \phi \, dx = - \int_{\Omega} f \partial_j \phi \, dx$$

for every  $\phi \in C_c^\infty(\Omega)$ . Similarly,  $f$  has weak derivative  $\partial^\alpha f = g$  if

$$\int_{\Omega} g \phi \, dx = (-1)^{|\alpha|} \int_{\Omega} f \partial^\alpha \phi \, dx.$$

The following lemma is fundamental.

**Lemma 2.20** *When they exist, weak derivatives are unique.*

*Proof* It suffices to show that if  $g \in L^1_{\text{loc}}(\Omega)$  and

$$\int_{\Omega} g \phi \, dx = 0$$

for every  $\phi \in C_c^\infty(\Omega)$  then  $g = 0$  almost everywhere.

Fix  $\Omega' \subset\subset \Omega$  and  $\epsilon > 0$ ; then  $g \in L^1(\Omega')$  and we can approximate  $g$  in  $L^1(\Omega')$  by a smooth function  $\tilde{g} \in C_c^\infty(\Omega')$  with

$$\|g - \tilde{g}\|_{L^1(\Omega')} < \epsilon/3.$$

Write  $\psi = \text{sgn}(\tilde{g})$ ; then  $|\psi| \leq 1$ , and so  $\psi \in L^1(\Omega')$ ; since  $\tilde{g}$  has compact

support in  $\Omega'$ , so does  $\psi$ . Approximate  $\psi$  by its mollification  $\psi_h$ , with  $h$  chosen small enough that  $\text{supp}(\psi_h) \subset \Omega'$  and

$$\|\psi - \psi_h\|_{L^1} < \frac{\epsilon}{3\|\tilde{g}\|_\infty}.$$

Note that

$$|\psi_h(x)| = \left| \int \rho_h(y-x)\psi(y) dy \right| \leq \int \rho_h(y-x)|\psi(y)| dy \leq \int \rho_h(y-x) dy = 1.$$

Then – noting that the support of  $\psi_h$  is a subset of  $\Omega'$ , and that  $\tilde{g}\psi = |\tilde{g}|$ ,

$$\begin{aligned} 0 &= \int_{\Omega} g\psi_h = \int_{\Omega'} g\psi_h \\ &= \int_{\Omega'} \tilde{g}\psi - \tilde{g}(\psi - \psi_h) + (g - \tilde{g})\psi_h \\ &\geq \int_{\Omega'} |\tilde{g}| - \int_{\Omega'} |\tilde{g}||\psi - \psi_h| - \int_{\Omega'} |g - \tilde{g}||\psi_h| \\ &= \|\tilde{g}\|_{L^1} - \|\tilde{g}\|_\infty \|\psi - \psi_h\|_{L^1} - \|g - \tilde{g}\|_{L^1} \\ &\geq \|g\|_{L^1} - \|g - \tilde{g}\|_{L^1} - \frac{\epsilon}{3} - \frac{\epsilon}{3}, \end{aligned}$$

and so for any  $\epsilon > 0$ ,

$$\|g\|_{L^1(\Omega')} \leq \epsilon$$

i.e.  $\|g\|_{L^1(\Omega')} = 0$ ,  $g = 0$  almost everywhere in  $\Omega'$ ; since  $\Omega' \subset\subset \Omega$  was arbitrary, this shows that  $g = 0$  almost everywhere in  $\Omega$ .  $\square$

As an example, consider first the function

$$f_1(x) = \begin{cases} x, & 0 < x \leq 1 \\ 1 & 1 < x < 2. \end{cases}$$

Then for any  $\phi \in C_c^\infty(0, 2)$

$$\begin{aligned} - \int_0^2 f_1(x)\phi'(x) dx &= - \int_0^1 x\phi'(x) dx - \int_1^2 \phi'(x) dx \\ &= - [x\phi(x)]_0^1 + \int_0^1 \phi(x) dx - \phi(2) + \phi(1) \\ &= \int_0^1 \phi(x) dx, \end{aligned}$$

since  $\phi(2) = 0$  (as  $\phi \in C_c^\infty(0, 2)$ ). So  $f_1$  has weak derivative

$$v(x) = \begin{cases} 1, & 0 < x \leq 1 \\ 0 & 1 < x < 2. \end{cases}$$

However, if we perform the same calculation starting with the function

$$f_2(x) = \begin{cases} x, & 0 < x \leq 1 \\ 2 & 1 < x < 2 \end{cases}$$

we end up with

$$\begin{aligned} - \int_0^2 f_2(x) \phi'(x) \, dx &= - \int_0^1 x \phi'(x) \, dx - 2 \int_1^2 \phi'(x) \, dx \\ &= - [x\phi(x)]_0^1 + \int_0^1 \phi(x) \, dx - 2\phi(2) + 2\phi(1) \\ &= \int_0^1 \phi(x) \, dx + \phi(1), \end{aligned}$$

There is no  $v \in L_{\text{loc}}^1(0, 2)$  such that

$$\int_0^2 v(x) \phi(x) \, dx = \phi(1).$$

Indeed, choose a sequence of  $\phi_n \in C_c^\infty(0, 2)$  such that  $0 \leq \phi_n(x) \leq 1$ ,  $\phi_n(1) = 1$ , and  $\phi_n(x) \rightarrow 0$  for all  $x \neq 1$ . Then, using the DCT,

$$1 = \lim_{n \rightarrow \infty} \phi_n(1) = \lim_{n \rightarrow \infty} \int_0^2 v(x) \phi_n(x) \, dx = 0$$

(to use the DCT:  $v(x)\phi_n(x) \rightarrow 0$  almost everywhere,  $|v(x)\phi_n(x)| \leq |v(x)|$ , and  $v \in L_{\text{loc}}^1(0, 2)$ ).

## 2.8.2 Sobolev spaces

**Definition 2.21** *The Sobolev space  $W^{k,p}(\Omega)$  is defined as*

$$W^{k,p}(\Omega) = \{f : \partial^\alpha f \in L^p(\Omega), 0 \leq |\alpha| \leq k\},$$

with norm

$$\|f\|_{W^{k,p}} = \left( \sum_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_{L^p}^p \right)^{1/p}.$$

**Lemma 2.22**  *$W^{k,p}(\Omega)$  is a Banach space, and separable if  $1 \leq p < \infty$ .*

*Proof* If  $\{f_n\}$  is a Cauchy sequence in  $W^{k,p}(\Omega)$  then  $\{\partial^\alpha f_n\}$  is Cauchy in  $L^p(\Omega)$  for every multi-index  $\alpha$  with  $0 \leq |\alpha| \leq k$ . So for each such  $\alpha$ , there

exists an  $f_\alpha \in L^p$  such that  $\partial^\alpha f \rightarrow f_\alpha$  in  $L^p$ . It remains to show that  $\partial^\alpha f_0 = f_\alpha$ .

To do this, take  $\phi \in C_c^\infty(\Omega)$ ; then

$$\begin{aligned} \int_\Omega f_\alpha \phi &= \lim_{n \rightarrow \infty} \int_\Omega (\partial^\alpha f_n) \phi = - \lim_{n \rightarrow \infty} \int_\Omega f_n (\partial^\alpha \phi) = - \int_\Omega f (\partial^\alpha \phi) \\ &= \int_\Omega (\partial^\alpha f) \phi, \end{aligned}$$

i.e.  $f_\alpha = \partial^\alpha f$ , and  $f \in W^{k,p}(\Omega)$ . [To justify swapping the sum and the integral, note that if  $f_n \rightarrow f$  in  $L^p$  and  $g \in L^q$  then

$$\left| \int f_n g - \int f g \right| = \left| \int (f_n - f) g \right| \leq \|f_n - f\|_{L^p} \|g\|_{L^q},$$

i.e.  $\int f_n g \rightarrow \int f g$ .]

For separability, see the problems sheet. □

We write  $H^k(\Omega) = W^{k,2}(\Omega)$ . This is a Hilbert space with the inner product

$$(f, g)_{H^k} = \sum_{0 \leq |\alpha| \leq k} (\partial^\alpha f, \partial^\alpha g)_{L^2}.$$

This inner product gives rise to the  $H^k$ -norm defined by

$$\|f\|_{H^k}^2 = \sum_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_{L^2}^2$$

(as above).

### 2.8.3 Completions

Given any normed space  $(X, \|\cdot\|)$ , there is an abstract method of constructing its ‘completion’  $(\mathbb{X}, \|\cdot\|_{\mathbb{X}})$  – essentially, the ‘smallest’ complete normed space that contains  $(X, \|\cdot\|)$  (or, in fact, an isometrically isomorphic copy) as a subset. One would like to ‘add limits of Cauchy sequences to  $X$ ’, but if we only have  $X$  we don’t have these limits (unless they’re already in  $X$ ). So instead, elements of  $\mathbb{X}$  consist of equivalence classes of Cauchy sequences in  $X$  (where  $\underline{X} \sim \underline{Y}$  if  $\lim_{n \rightarrow \infty} \|X_n - Y_n\| = 0$ ). With some care – much of the work is in coping with the notation – you can show that the resulting space *is* complete, and that there’s a copy of  $X$  (consisting of (equivalence classes of) the constant sequences with  $X_j = x$  for every  $j$ ) inside  $\mathbb{X}$ .

However, the situation is much more straightforward when we already have a space  $(X, \|\cdot\|)$ , and we want to find the completion of a subspace *in*  $X$ . Suppose that  $Y$  is a subspace of  $X$ ; then the completion of  $Y$  in  $X$  is simply the closure of  $Y$  in  $X$ , i.e.

$$\bar{Y} = \{x \in X : x = \lim_{n \rightarrow \infty} y_n, \text{ for some } y_n \in Y\}.$$

In this case we really can ‘add the limit of any Cauchy sequence’, since if  $\{y_n\}$  is Cauchy we know that the limit exists and lies in  $X$ . It is simple to show that  $(Y, \|\cdot\|)$  is complete.

Note that if  $Y$  is dense in  $X$ , then the completion of  $Y$  in  $X$  is all of  $X$ , by definition; we have already seen, for example, that  $C_c^\infty(\Omega)$  is dense in  $L^2(\Omega)$ , and so the completion of  $C_c^\infty(\Omega)$  in  $L^2(\Omega)$  is  $L^2(\Omega)$ .

However,  $C_c^\infty(\Omega)$  is *not* dense in  $H^k(\Omega)$  if  $k \geq 1$ . Suppose that  $f$  is a smooth non-zero solution of

$$\sum_{|\alpha| \leq k} (-1)^{|\alpha|} \partial^{2\alpha} f = 0$$

on  $\Omega$ ; in particular  $f \in H^k$ . [Solve the equation using separation of variables on a parallelepiped containing  $\Omega$ , then restrict this to  $\Omega$ .]

Since  $f \neq 0$ ,  $\|f\|_{H^k}^2 \neq 0$ . Now suppose that there is a sequence  $\phi_j \in C_c^\infty(\Omega)$  such that  $\phi_j \rightarrow f$  in  $H^k$ . Then

$$\begin{aligned} (f, f)_{H^k} &= \lim_{j \rightarrow \infty} (f, \phi_j)_{H^k} \\ &= \lim_{j \rightarrow \infty} \sum_{|k| \leq \alpha} (\partial^\alpha f, \partial^\alpha \phi_j)_{L^2} \\ &= \lim_{j \rightarrow \infty} \sum_{|k| \leq \alpha} ((-1)^{|\alpha|} \partial^{2\alpha} f, \phi_j)_{L^2} \\ &= \lim_{j \rightarrow \infty} \left( \sum_{|k| \leq \alpha} (-1)^{|\alpha|} \partial^{2\alpha} f, \phi_j \right)_{L^2} \\ &= 0, \end{aligned}$$

a contradiction. [Note: since  $(f, \cdot)_{H^k}$  defines a bounded linear functional on  $H^k$ , this shows that there are non-zero bounded linear functionals on  $H^k$ , that nevertheless vanish on all test functions.]

The completion of  $C_c^\infty(\Omega)$  in  $H^k$  is very useful, and we denote this space

by  $H_0^k(\Omega)$ . Heuristically, this space consists of functions in  $H^k$  with  $\partial^\alpha f = 0$  on  $\partial\Omega$  for  $|\alpha| \leq k - 1$ .

In light of this the following version of the Poincaré inequality is perhaps unsurprising.

**Lemma 2.23** *Suppose that  $\Omega$  is bounded in one direction in a strip of width  $d$ . Then for every  $f \in H_0^1(\Omega)$ ,*

$$\|f\|_{L^2} \leq d\|\nabla f\|_{L^2}. \quad (2.10)$$

*Proof* We have already seen that if  $f \in C_c^\infty(\Omega)$  then the inequality holds. Now take  $f \in H_0^1(\Omega)$ , and write  $f = \lim_{n \rightarrow \infty} f_n$ , with  $f_n \in C_c^\infty(\Omega)$  and the limit taken in the  $H^1$  norm. This means in particular, that  $f_n \rightarrow f$  in  $L^2$  and  $\partial_j f_n \rightarrow \partial_j f$  in  $L^2$ ; so

$$\|f\|_{L^2} = \lim_{n \rightarrow \infty} \|f_n\|_{L^2}$$

and

$$\|\nabla f\|_{L^2}^2 := \sum_{j=1}^n \|\partial_j f\|_{L^2}^2 = \lim_{n \rightarrow \infty} \sum_{j=1}^n \|\partial_j f_n\|_{L^2}^2 =: \|\nabla f_n\|_{L^2}^2,$$

whence (2.10). □

The space  $H^{-k}(\Omega)$  is the dual space of  $H_0^k(\Omega)$  [i.e. the space of all bounded linear functionals on  $H_0^k(\Omega)$ ]. (We do not want to have non-zero elements of the dual that vanish on all test functions, as we did above.)

#### 2.8.4 Density results for $H^k(\Omega)$

However,  $C^\infty(\Omega)$  is dense in  $H^k$ , which will follow from the following pleasing result – mollification and derivatives commute.

**Lemma 2.24** *Suppose that  $f \in L_{\text{loc}}^1(\Omega)$  and that the weak derivative  $\partial^\alpha f$  exists. Then provided that  $h < \text{dist}(x, \partial\Omega)$ ,*

$$\partial^\alpha f_h(x) = (\partial^\alpha f)_h(x).$$

*Proof* Using the definition of  $f_h$  and differentiating under the integral sign, we have, integrating by parts

$$\begin{aligned}\partial^\alpha f_h(x) &= \int_{\Omega} \partial_x^\alpha \rho_h(x-y) f(y) \, dy \\ &= (-1)^{|\alpha|} \int_{\Omega} \partial_y^\alpha \rho_h(x-y) f(y) \, dy \\ &= \int_{\Omega} \rho_h(x-y) \partial^\alpha f(y) \, dy \\ &= (\partial^\alpha f)_h.\end{aligned}$$

□

We define the space  $H_{\text{loc}}^k(\Omega)$  in the obvious way:  $f \in H_{\text{loc}}^k(\Omega)$  if  $f \in H^k(\Omega')$  for every  $\Omega' \subset\subset \Omega$ , and  $f_n \rightarrow f$  in  $H_{\text{loc}}^k(\Omega)$  if  $f_n \rightarrow f$  in  $H^k(\Omega')$  for every  $\Omega' \subset\subset \Omega$ .

**Corollary 2.25** *If  $f \in H_{\text{loc}}^k(\Omega)$  then  $f_h \in C^\infty(\Omega) \cap H_{\text{loc}}^k(\Omega)$  and  $f_h \rightarrow f$  in  $H_{\text{loc}}^k(\Omega)$ .*

Given this, we require one other technical device, and the proof that  $C^\infty(\Omega)$  is dense in  $H^k(\Omega)$  is then ‘easy’. This device is known as a ‘partition of unity’.

**Theorem 2.26** *Let  $\Omega$  be an open subset of  $\mathbb{R}^n$ , and  $\{U_j\}$  a countable collection of bounded open subsets of  $\mathbb{R}^n$  that cover  $\Omega$ , with the additional property that every compact subset of  $\Omega$  intersects only finitely many of the  $\{U_j\}$  (the cover is ‘locally finite’). Then there exists a partition of unity subordinate to the covering  $\{U_j\}$ , that is a set of functions  $\{\psi_j\} \in C_c^\infty(\mathbb{R}^n)$  such that*

- (i)  $0 \leq \psi_j \leq 1$ ,
- (ii)  $\text{supp} \psi_j \subset U_j$ , and
- (iii)  $\sum_{j=1}^{\infty} \psi_j = 1$  on a neighbourhood of  $\Omega$ .

For a proof see Theorem 5.14 in [JCR].

**Theorem 2.27**  *$C^\infty(\Omega) \cap H^k(\Omega)$  is dense in  $H^k(\Omega)$ .*

*Proof* Set

$$\Omega_j = \{x \in \Omega : \text{dist}(x, \partial\Omega) > 1/j \quad \text{and} \quad |x| < j\}$$

Then  $\Omega_j \subset\subset \Omega_{j+1}$  and  $\Omega = \cup_{j=1}^{\infty} \Omega_j$ . Let  $\{\psi_j\}$  be a partition of unity subordinate to the covering  $\Omega_{j+1} \setminus \overline{\Omega_j}$ .

Fix  $f \in H^k(\Omega)$  and  $\epsilon > 0$ . For each  $j$  choose

$$h_j < \text{dist}(\Omega_j, \partial\Omega_{j+1})$$

such that

$$\|(\psi_j u)_{h_j} - \psi_j u\|_{H^k(\Omega)} \leq 2^{-j} \epsilon.$$

By the choice of  $h_j$  and  $\psi_j$ , only a finite number of the functions  $(\psi_j u)_{h_j}$  are non-zero on any compact subdomain of  $\Omega$ ; it follows that the function

$$\tilde{u} = \sum_{j=1}^{\infty} (\psi_j u)_{h_j}$$

is in  $C^\infty(\Omega)$ . Furthermore

$$\begin{aligned} \|u - \tilde{u}\|_{H^k(\Omega)} &= \sum_{j=1}^{\infty} \left\| \left( \sum_{j=1}^{\infty} \psi_j \right) u - \tilde{u} \right\| \\ &\leq \sum_{j=1}^{\infty} \|(\psi_j u)_{h_j} - \psi_j u\| \\ &\leq \sum_{j=1}^{\infty} 2^{-j} \epsilon = \epsilon. \end{aligned}$$

□

Note that there are minimal assumptions here on  $\Omega$ .

We can use this result to prove the following ‘approximation up to the boundary’ in a half space.

**Theorem 2.28** *The space  $C^\infty(\overline{\mathbb{R}_+^n})$  is dense in  $H^k(\mathbb{R}_+^n)$ .*

We denote by  $\tau_h f$  the ‘shifted function’  $\tau_h f(x) = f(x', x_n + h)$  (note that this shifts the values of the functions ‘downwards’).

*Proof* Take  $f \in H^k(\mathbb{R}_+^n)$ . Then we know that there exist  $f_n \in C^\infty(\mathbb{R}_+^n)$  such that  $f_n \rightarrow f$  in  $H^k(\mathbb{R}_+^n)$ . Now consider the shifted functions

$$\tau_{1/n} f_n(x) := f_n(x', x_n + \frac{1}{n}),$$

where  $x' = (x_1, \dots, x_{n-1})$ . Clearly  $\tau_{1/n}f_n$  is defined for all  $x_n > 1/n$ , and

$$\partial^\alpha(\tau_{1/n}f_n) = \tau_{1/n}\partial^\alpha f_n.$$

Denote by  $\tilde{f}_n$  the restriction of  $\tau_{1/n}f_n$  to  $\overline{\mathbb{R}_+^n}$ . Then  $\tilde{f}_n \in H^k(\mathbb{R}_+^n) \cap C^\infty(\overline{\mathbb{R}_+^n})$ , and for any  $\alpha$  with  $0 \leq |\alpha| \leq k$ ,

$$\|\partial^\alpha(f - \tilde{f}_n)\|_{L^2(\mathbb{R}_+^n)} \leq \|\partial^\alpha(f - f_n)\|_{L^2(\mathbb{R}_+^n)} + \|\partial^\alpha(f_n - \tilde{f}_n)\|_{L^2(\mathbb{R}_+^n)}.$$

The first term tends to zero by our choice of  $f_n$ ; the second tends to zero since for any  $g \in L^2(\mathbb{R}_+^n)$ ,

$$\|\tau_h g - g\|_{L^2(\mathbb{R}_+^n)} \rightarrow 0 \quad \text{as} \quad h \rightarrow 0.$$

(You can prove this using step functions, see examples.) □

### 2.8.5 Extension theorems

First, extending functions in  $H^k(\mathbb{R}_+^n)$ , where

$$\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_n > 0\}.$$

We show that if  $u \in H^k(\mathbb{R}_+^n)$  then there is a linear operator  $E_k : H^k(\mathbb{R}_+^n) \rightarrow H^k(\mathbb{R}^n)$  such that

$$\|E_k[u]\|_{H^k(\mathbb{R}^n)} \leq c_k \|u\|_{H^k(\mathbb{R}_+^n)}.$$

We then use a change of coordinates to treat sufficiently smooth bounded domains.

**Theorem 2.29** *Let  $k \geq 0$  be an integer. Then there exists a bounded linear mapping  $E_k : H^k(\mathbb{R}_+^n) \rightarrow H^k(\mathbb{R}^n)$  such that*

$$E[f]|_{\mathbb{R}_+^n} = f \quad \text{for all} \quad f \in H^k(\mathbb{R}_+^n)$$

and

$$\|E_k[f]\|_{H^j(\mathbb{R}^n)} \leq c_k \|f\|_{H^j(\mathbb{R}_+^n)} \quad \text{for all} \quad j = 1, \dots, k.$$

*Proof* First we consider how to extend functions in  $H^k(\overline{\mathbb{R}_+^n}) \cap C^k(\mathbb{R}_+^n)$ , and then we use a density argument to treat any  $f \in H^k(\mathbb{R}_+^n)$ .

To make  $f$  continuous across  $x_n = 0$  we could simply define

$$f(x', x_n) = f(x', -x_n) \quad \text{for all} \quad x_n > 0,$$

i.e. ‘reflect’ in  $x_n = 0$ . But this would cause problems with the first derivative; to get round these we could try

$$f(x', x_n) = 3f(x', -x_n) - 2f(x', -2x_n),$$

which now matches by  $f$  and  $f'$  across  $x_n = 0$ . To match derivatives up to order  $k$ , try a combination of the form

$$f(x', x_n) = \sum_{j=1}^{k+1} a_j f(x', -jx_n).$$

For the first  $k$  derivatives to agree requires

$$\sum_{j=1}^{k+1} (-j)^i a_j = 1 \quad \text{for all } i = 0, \dots, k.$$

This equation for  $\{a_j\}$  always has a solution: to see this, consider the matrix equation

$$J\underline{a} = \underline{1},$$

where  $J$  is the  $(k+1) \times (k+1)$  matrix with  $J_{ij} = (-j)^{i-1}$ . Now,  $J$  is invertible: if not there must be some non-zero  $\underline{d}$  such that  $J\underline{d} = 0$ . But this implies that the degree  $k$  polynomial

$$\sum_{i=0}^k d_{i+1} x^i$$

has  $k+1$  roots,  $-1, \dots, -(k+1)$ .

That’s it. □

Let  $U$  and  $V$  be open subsets of  $\mathbb{R}^n$ . A function  $\Phi : U \rightarrow V$  is a  $C^k$ -diffeomorphism if

- (i)  $\Phi$  and  $\Phi^{-1}$ , and their derivatives of order up to  $k$ , are bounded and continuous on  $U$  and  $V$  respectively, and
- (ii) there are positive constants  $k_1$  and  $k_2$  such that

$$k_1 \leq |\det \nabla \Phi(x)| \leq k_2$$

for all  $x \in U$ .

Suppose that we have a function  $f : U \rightarrow \mathbb{R}$ . If we change coordinates we produce a new function  $f^* : V \rightarrow \mathbb{R}$ , given by

$$f^*(y) = f(\Phi^{-1}(y)).$$

This is the ‘pullback of  $f$  under  $\Phi^{-1}$ ’,  $(\Phi^{-1})^*f = f \circ \Phi^{-1}$ . Note that  $\Phi^{-1}$  acts on points in  $V$ ;  $(\Phi^{-1})^*$  acts on functions defined on  $U$ . Similarly, a function defined on  $V$  gives rise to a function defined on  $U$  via the pullback operator  $\Phi^*$ ,

$$(\Phi^*g)(x) = g(\Phi(x)).$$

The following lemma – whose proof we omit – shows that if  $f \in H^k(U)$  then  $(\Phi^{-1})^*f \in H^k(V)$ , and similarly if  $g \in H^k(V)$  then  $\Phi^*g \in H^k(U)$ .

**Lemma 2.30** *If  $\Phi : U \rightarrow V$  is a  $C^k$ -diffeomorphism then  $(\Phi^{-1})^*$  and  $\Phi^*$  are bounded linear maps from  $H^k(U)$  to  $H^k(V)$ , and vice versa, i.e. for  $f \in H^k(U)$  and  $g \in H^k(V)$ ,*

$$\|(\Phi^{-1})^*f\|_{H^k(V)} \leq c\|f\|_{H^k(U)} \quad \text{and} \quad \|\Phi^*g\|_{H^k(U)} \leq c\|g\|_{H^k(V)}.$$

Now we suppose that  $\Omega$  is a  $C^k$  domain: this means that at each point  $x_0 \in \partial\Omega$  there exists an  $\epsilon > 0$  and a  $C^k$  diffeomorphism  $\Phi$  of  $B(x_0, \epsilon)$  onto a subset  $B$  of  $\mathbb{R}^n$  such that

- (i)  $\Phi(x_0) = 0$ ,
- (ii)  $\Phi(B(x_0, \epsilon) \cap \Omega) \subset \mathbb{R}_+^n$ , and
- (iii)  $\Phi(B(x_0, \epsilon) \cap \partial\Omega) \subset \partial\mathbb{R}_+^n$ .

**Theorem 2.31** *If  $\Omega$  is a bounded  $C^k$  domain, then for each open set  $\Omega^*$  with  $\Omega \subset\subset \Omega^*$  there exists a bounded linear extension operator  $E$  such that if  $f \in H^k(\Omega)$  then  $E[f] \in H_0^k(\Omega^*)$  and*

$$\|E[f]\|_{H^k(\Omega^*)} \leq C_{k,\Omega^*}\|f\|_{H^k(\Omega)}. \quad (2.11)$$

(In fact (2.11) holds for each  $H^j$  with  $0 \leq j \leq k$ .)

*Proof* At each  $x \in \partial\Omega$  find an  $\epsilon_x$  such that there exists a  $C^k$  diffeomorphism of  $B(x, \epsilon_x)$  onto a subset of  $\mathbb{R}^n$  as above, and such that  $B(x, \epsilon_x) \subset\subset \Omega^*$ . Since  $\Omega$  is bounded so is  $\partial\Omega$ , which is therefore compact: choose a finite collection of the balls  $B(x, \epsilon_x)$ , call them  $\{B_j\}_{j=1}^m$ , that cover an open

neighbourhood of  $\partial\Omega$ , and corresponding functions  $\Phi_j$ . Write  $U_j = \Omega \cap B_j$  and  $V_j = \Phi_j(\Omega_j)$

Let  $\{\psi_j\}$  be a partition of unity subordinate to the  $\{B_j\}$ . Then

$$\sum_{j=1}^m \psi_j f$$

is equal to  $f$  in a neighbourhood of  $\partial\Omega$ , and the function defined within  $\Omega$  as

$$\tilde{f} := f - \sum_{j=1}^m \psi_j f$$

has compact support within  $\Omega$ . Clearly within  $\Omega$  we have

$$f = \tilde{f} + \sum_{j=1}^m \psi_j f.$$

For each  $j$  consider the function  $f_j^* \in H^k(\mathbb{R}_+^n)$  defined by

$$f_j^* = (\Phi_j^{-1})^*(\psi_j f),$$

which is zero outside  $V_j$ . By Lemma 2.30,

$$\|f_j^*\|_{H^k(\mathbb{R}_+^n)} \leq c_j \|\psi_j f\|_{H^k(U_j)},$$

and since for any  $\psi \in C_c^\infty(\mathbb{R}^n)$ ,  $U \subset \mathbb{R}^n$  and  $f \in H^k(U)$  we have

$$\|\psi f\|_{H^k(U)} \leq C(\psi) \|f\|_{H^k(U)},$$

it follows that

$$\|f_j^*\|_{H^k(\mathbb{R}_+^n)} \leq c'_j \|f\|_{H^k(U_j)} \leq c'_j \|f\|_{H^k(\Omega)}.$$

Now, we know that the function  $f_j^*$  can be extended to a function  $E_k[f_j^*] = f_j \in H^k(\mathbb{R}^n)$ , such that

$$\|f_j\|_{H^k(\mathbb{R}^n)} \leq C_k \|f_j^*\|_{H^k(\mathbb{R}_+^n)} \leq c''_j \|f\|_{H^k(\Omega)}.$$

Now multiply  $f_j$  by a  $C_c^\infty(\mathbb{R}^n)$  cutoff function  $\theta_j$ , that is equal to 1 on  $\Phi_j(\text{supp}(\psi_j))$  and has compact support within  $\Phi_j(B_j)$ . Then the pullback function

$$\Phi_j^*(\theta_j f_j)$$

is equal to  $\psi_j u$  on  $U_j$ , has compact support within  $B_j$  (and so within  $\Omega^*$ ), and satisfies

$$\|\Phi_j^*(\theta_j f_j)\|_{H^k(\Omega^*)} \leq C_j \|f\|_{H^k(\Omega)}. \quad (2.12)$$

Since  $\tilde{f}$  has compact support within  $\Omega$ , its extension by zero – which we still write as  $\tilde{f}$  is in  $H^k(\Omega)$ .

So now consider the composite function

$$E[f] = \tilde{f} + \sum_{j=1}^m \Phi_j^*(\theta_j \tilde{f}_j).$$

This extends  $u$  to a function in  $H_0^k(\Omega^*)$ , and on summing the estimates in (2.12) has

$$\|E[f]\|_{H^k(\Omega^*)} \leq \mathfrak{C}_k \|f\|_{H^k(\Omega)}.$$

□

Now we can prove...

**Theorem 2.32** *If  $\Omega$  is of class  $C^k$  and  $\partial\Omega$  is compact, then  $C^\infty(\bar{\Omega})$  is dense in  $H^k(\Omega)$ .*

*Proof* Take  $\Omega^*$  such that  $\Omega \subset\subset \Omega^*$ , and given  $f \in H^k(\Omega)$  extend to a function  $E[f] \in H_0^k(\Omega^*)$ . Then there exists a sequence of functions  $g_n \in C^\infty(\Omega^*) \cap H^k(\Omega^*)$  such that  $g_n \rightarrow E[f]$  in  $H^k(\Omega^*)$ . Set  $f_n = g_n|_\Omega$ ; then  $f_n \in C^\infty(\bar{\Omega})$  and

$$\|f_n - f\|_{H^k(\Omega)} \leq \|g_n - E[f]\|_{H^k(\Omega^*)} \rightarrow 0$$

as  $n \rightarrow \infty$ .

□

## 2.9 The Fourier transform and Sobolev spaces

Before proving various Sobolev embedding theorems we give a (very) brief treatment of the Fourier transform (FT).

Let  $\mathcal{S}(\mathbb{R}^n)$  be the class of all complex-valued  $C^\infty$  functions on  $\mathbb{R}^n$  such that  $|x|^k |\partial^\alpha f(x)|$  is bounded for every  $k \in \mathbb{N}$  and every multi-index  $\alpha$ .

The Fourier transform of any  $f \in \mathcal{S}(\mathbb{R}^n)$  is defined as

$$\hat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) dx.$$

It is relatively easy to show that if  $f \in \mathcal{S}(\mathbb{R}^n)$  then  $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$ .

If  $g \in \mathcal{S}(\mathbb{R}^n)$  then there exists a unique  $f \in \mathcal{S}(\mathbb{R}^n)$  such that  $\hat{f} = g$ , and

$$f(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\xi \cdot x} g(x) \, dx,$$

i.e.  $f = (\check{g})^\wedge$ , where  $\check{g}(x) = g(-x)$ .

The FT is particularly useful because it ‘converts’ derivatives into ‘multipliers’:

$$(\partial^\alpha f)^\wedge = (i\xi)^\alpha \hat{f}.$$

Also, the FT preserves the  $L^2$  inner product: if  $f, g \in \mathcal{S}(\mathbb{R}^n)$  then

$$(\hat{f}, \hat{g}) = (f, g);$$

in particular

$$\|\hat{f}\|_{L^2} = \|f\|_{L^2}.$$

Since the FT is linear, this enables us to define the FT for functions  $f \in L^2(\mathbb{R}^n)$  using a limit procedure. Take  $f \in L^2(\mathbb{R}^n)$ ; then since  $\mathcal{S}(\mathbb{R}^n)$  is dense in  $L^2(\mathbb{R}^n)$  (we know that  $C_c^\infty(\mathbb{R}^n)$  functions are dense in  $L^2(\mathbb{R}^n)$  and these are a subset of functions in  $\mathcal{S}(\mathbb{R}^n)$ ), we can find  $f_n \in \mathcal{S}(\mathbb{R}^n)$  such that  $f = \lim_{n \rightarrow \infty} f_n$  (where the limit is taken in  $L^2$ ), and then define

$$\hat{f} = \lim_{n \rightarrow \infty} \hat{f}_n,$$

where again the limit is taken in  $L^2(\mathbb{R}^n)$ . [So if  $f \in L^2$ ,  $\hat{f} \in L^2$ .]

If  $f \in H^k(\mathbb{R}^n)$ , then the derivatives  $\partial^\alpha f$  are in  $L^2(\mathbb{R}^n)$  for  $0 \leq |\alpha| \leq k$ ; it follows that

$$(i\xi)^\alpha \hat{f} \in L^2(\mathbb{R}^n) \quad \text{for all} \quad 0 \leq |\alpha| \leq k,$$

which we can combine to give a characterisation of  $H^k$  in terms of the FT,

$$H^k(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : (1 + |\xi|^2)^{k/2} \hat{f} \in L^2(\mathbb{R}^n)\}, \quad (2.13)$$

and it is straightforward to show that the norm

$$\|(1 + |\xi|^2)^{k/2} \hat{f}\|_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} (1 + |\xi|^2)^k |\hat{f}(\xi)|^2 \, d\xi$$

is equivalent to the norm in  $H^k(\mathbb{R}^n)$ . [Note that there is no reason why  $k$  should be an integer in (2.13), and this gives a way to define Sobolev spaces with non-integral exponents.]

**Theorem 2.33** *Let  $f \in H^k(\mathbb{R}^n)$  with  $k > n/2$ . Then there exists a constant  $C_{k,n}$  such that*

$$\|f\|_\infty \leq C_{k,n} \|f\|_{H^k(\mathbb{R}^n)},$$

*and  $f$  is equal almost everywhere to a function in  $C^0(\mathbb{R}^n)$ .*

Actually the argument here shows also that the FT maps  $L^1$  to  $L^\infty$ .

*Proof* First take  $f \in \mathcal{S}(\mathbb{R}^n) \cap H^k(\mathbb{R}^n)$ ; then

$$\begin{aligned} \|f\|_\infty &= \sup_{x \in \mathbb{R}^n} (2\pi)^{-n/2} \left| \int_{\mathbb{R}^n} e^{i\xi \cdot x} \hat{f}(\xi) \, d\xi \right| \\ &\leq (2\pi)^{-n/2} \int |\hat{f}(\xi)| \, d\xi \\ &= (2\pi)^{-n/2} \int \frac{1}{(1+|\xi|^2)^{k/2}} (1+|\xi|^2)^{k/2} |\hat{f}(\xi)| \, d\xi \\ &\leq (2\pi)^{-n/2} \left( \int \frac{1}{(1+|\xi|^2)^k} \, d\xi \right)^{1/2} \left( \int (1+|\xi|^2)^k |\hat{f}(\xi)|^2 \, d\xi \right)^{1/2} \\ &= C_{k,n} \|f\|_{H^k(\mathbb{R}^n)}, \end{aligned}$$

using the fact (see exercises) that

$$\int_{\mathbb{R}^n} \frac{1}{(1+|\xi|^2)^k} \, d\xi < \infty$$

if and only if  $k > n/2$ . The same inequality follows for  $f \in H^k(\mathbb{R}^n)$  by taking limits. But also, this shows that  $f$  is the  $L^\infty$  limit of a sequence of continuous functions; it must therefore be equal almost everywhere to a continuous function, and the result follows.  $\square$

**Corollary 2.34** *If  $f \in H^k(\Omega)$  with  $k > n/2 + s$  then  $f \in C^s(\bar{\Omega})$  with  $\|f\|_{C^s} \leq \|f\|_{H^k}$ .*

*Proof* Take  $f \in H^k(\Omega)$ : then there is an extension of  $f$  to  $\mathbb{R}^n$ ,  $E[f]$ , such that

$$\|E[f]\|_{H^k(\mathbb{R}^n)} \leq C_k^E \|f\|_{H^k(\Omega)}.$$

By the previous theorem,  $E[f] \in C^0(\mathbb{R}^n)$  with

$$\|E[f]\|_{C^0(\mathbb{R}^n)} \leq C_{k,n} \|E[f]\|_{H^k(\mathbb{R}^n)} \leq C_{k,n} C_k^E \|f\|_{H^k(\Omega)}.$$

Clearly, therefore,  $f \in C^0(\bar{\Omega})$  with

$$\|f\|_{C^0(\bar{\Omega})} \leq C_{\Omega,k,n} \|f\|_{H^k(\Omega)}.$$

The results for higher derivatives follow by considering each  $\partial^\alpha f$  in turn.  $\square$

Actually we can do better, and show that  $f$  must be Hölder continuous.

**Theorem 2.35** *If  $f \in H^k(\mathbb{R}^n)$  with  $n/2 < k < n/2 + 1$  then there is a constant  $C$  such that*

$$|f(x) - f(y)| \leq C_k \|f\|_{H^k} |x - y|^{k-(n/2)} \quad \text{for all } x, y \in \mathbb{R}^n.$$

*Proof* See examples.  $\square$

### 2.9.1 Compact embeddings

We say that space  $X$  is *compactly embedded* in  $Y$  if a bounded set in  $X$  is a precompact subset of  $Y$ , i.e. if any bounded sequence that is bounded in the norm of  $X$  has a subsequence that converges in the norm of  $Y$ . The prototype such theorem for us is the Arzelà-Ascoli Theorem – for a proof see the examples.

**Theorem 2.36 (Arzelà-Ascoli Theorem)** *Let  $K \subset \mathbb{R}^n$  be compact and let  $\{f_n\}$  be a sequence in  $C^0(K)$  that is*

- (i) *bounded, i.e.  $\|f_n\|_\infty \leq M$  for some  $M > 0$ , and*
- (ii) *equicontinuous, i.e. for every  $\epsilon > 0$ , there exists a  $\delta > 0$  such that*

$$|x - y| < \delta \quad \Rightarrow \quad |f_n(x) - f_n(y)| < \epsilon$$

*for every  $x \in K$ , and for every  $n \in \mathbb{N}$ .*

*Then  $\{f_n\}$  has a subsequence that converges uniformly on  $K$ .*

Using this result, the Hölder continuity above implies that if  $\Omega$  is bounded, the embedding of  $H^k(\Omega)$  into  $C^0(\bar{\Omega})$  is compact, i.e. a bounded sequence in  $H^k(\Omega)$  has a subsequence that converges uniformly on  $\bar{\Omega}$ .

The following result is similar, but arguably more important.

**Theorem 2.37** *If  $\Omega$  is bounded and  $\partial\Omega$  is  $C^k$  then  $H^{k+1}(\Omega)$  is compactly embedded in  $H^k(\Omega)$ .*

*Proof* Choose  $\Omega^*$  with  $\Omega \subset\subset \Omega^*$ . Take a sequence  $\{f_n\}$  that is bounded in  $H^{k+1}(\Omega)$ ,  $\|f_n\|_{H^{k+1}(\Omega)} \leq M$ , and extend to  $E[f_n]$ , bounded in  $H^{k+1}(\Omega^*)$  with  $\|E[f_n]\|_{H^{k+1}(\mathbb{R}^n)} \leq M'$ . For any  $g \in H^{k+1}(\Omega)$ , let

$$g_R = \mathcal{F}^{-1}[\hat{g}\chi_{|\xi| \leq R}];$$

then there exists a constant  $C$  such that

$$\|g - g_R\|_{H^k} \leq \frac{C}{R} \|g\|_{H^{k+1}};$$

Indeed,

$$\begin{aligned} \|g - g_R\|_{H^k}^2 &\leq C \int_{|\xi| \geq R} (1 + |\xi|^2)^k |\hat{g}(\xi)|^2 d\xi \\ &= \frac{C}{1 + R^2} \int_{|\xi| \geq R} (1 + |\xi|^2)^{k+1} |\hat{g}(\xi)|^2 d\xi \\ &\leq \frac{C}{1 + R^2} \|g\|_{H^{k+1}}^2. \end{aligned}$$

Similarly,  $g_R \in H^l(\mathbb{R}^n)$  for every  $l \in \mathbb{N}$ ;

$$\begin{aligned} \|g_R\|_{H^l(\mathbb{R}^n)} &\leq C \int_{|\xi| \leq R} (1 + |\xi|^2)^l |\hat{g}(\xi)|^2 d\xi \\ &\leq C(1 + R^2)^{l-(k+1)} \int_{|\xi| \leq R} (1 + |\xi|^2)^{k+1} |\hat{g}(\xi)|^2 d\xi \\ &= C_{R,l} \|g\|_{H^{k+1}}^2. \end{aligned}$$

So

$$\|g_R\|_{C_b^r(\mathbb{R}^n)} \leq C_{R,r} \|g\|_{H^{k+1}}.$$

It follows that if  $U$  is any bounded open set containing  $\Omega^*$ , then  $(E[f_n])_R$  has a subsequence that converges in  $C_b^k(\bar{U})$ , and hence in  $H^k(U)$ . applying a standard diagonal argument, we can obtain a subsequence such that  $(E[f_{n_j}])^m$  converges in  $H^k(D)$  for every  $m$ .

From this it follows that  $E[f_{n_j}]$  is Cauchy in  $H^k(U)$ : given  $\epsilon > 0$ , first choose  $m$  such that

$$\|E[f_i] - E[f_i]_m\|_{H^k} \leq \frac{\sqrt{C}}{m} \|E[f_i]\|_{H^{k+1}} \leq \frac{\sqrt{C}M'}{m} < \epsilon/3.$$

Then choose  $N$  such that

$$\|E[f_{n_i}]_m - E[f_{n_j}]_m\|_{H^k(U)} \leq \epsilon/3 \quad \text{for all } i, j \geq N.$$

It follows that

$$\begin{aligned} \|E[f_{n_i}] - E[f_{n_j}]\|_{H^k(U)} &\leq \|E[f_{n_i}] - E[f_{n_i}]_m\|_{H^k(U)} + \|E[f_{n_i}]_m - E[f_{n_j}]_m\|_{H^k(U)} \\ &\quad + \|E[f_{n_j}]_m - E[f_{n_j}]\|_{H^k(U)} \\ &\leq \epsilon \quad \text{for all } i, j \geq N, \end{aligned}$$

and so  $\{E[f_{n_i}]\}$  is Cauchy in  $H^k(U)$ . It follows that  $f_{n_i}$  is Cauchy in  $H^k(\Omega)$ , and since  $H^k(\Omega)$  is complete,  $\{f_{n_i}\}$  converges.  $\square$

### 2.9.2 Sobolev spaces and $L^p$ spaces

What happens if  $f \in H^k$  with  $k < n/2$ ? To consider this, we need the fact that the Fourier transform maps  $L^p$  into  $L^q$ , with  $(p, q)$  conjugate, such that

$$\|f\|_{L^p} \leq C_p \|\hat{f}\|_{L^q}. \quad (2.14)$$

This is certainly not obvious; it follows ‘by interpolation’ from the fact that it maps  $L^2$  to  $L^2$  and  $L^1$  to  $L^\infty$ .

**Theorem 2.38** *If  $f \in H^k(\mathbb{R}^n)$  with  $k \leq n/2$  then  $f \in L^p(\mathbb{R}^n)$  for any  $p \in [2, \frac{2n}{n-2k})$ .*

*Proof* Using (2.14) we have (for  $f \in H^k \cap \mathcal{S}(\mathbb{R}^n)$ )

$$\begin{aligned} \|f\|_{L^p}^q &\leq C_p^q \|\hat{f}\|_{L^q}^q \\ &= C_p^q \int |\hat{f}(\xi)|^q d\xi \\ &= C_p^q \int \frac{1}{(1 + |\xi|^2)^{kq/2}} (1 + |\xi|^2)^{kq/2} |\hat{f}(\xi)|^q d\xi \\ &\leq C_p^q \left( \int \frac{1}{(1 + |\xi|^2)^{kq/2}} (1 + |\xi|^2)^{2(2-q)/2} d\xi \right)^{(2-q)/2} \left( \int (1 + |\xi|^2)^k |\hat{f}(\xi)|^2 d\xi \right)^{q/2} \\ &= C_p^q \left( \int \frac{1}{(1 + |\xi|^2)^{kq/(2-q)}} d\xi \right)^{(2-q)/2} \|f\|_{H^k(\mathbb{R}^n)}^q \\ &\leq C_{p,k} \|f\|_{H^k(\mathbb{R}^n)}^q. \end{aligned}$$

The condition

$$\frac{kq}{2-q} > \frac{n}{2} \Leftrightarrow \frac{1}{q} < \frac{2k+n}{2n}$$

ensures that the first integral is finite; this translates to

$$\frac{1}{p} > 1 - \frac{2k+n}{2n} \Leftrightarrow p < \frac{2n}{n-2k}.$$

□

Note that if  $k = n/2$  this shows that  $f \in L^p$  for any  $2 \leq p < \infty$ . When  $k < n/2$  the result is also true for the endpoint  $p = 2n/(n-2k)$ , which is the most useful; this requires the ‘death by Young’s inequality’ approach.

**Corollary 2.39** *If  $f \in H^k(\Omega)$  with  $k \leq n/2$  then  $f \in L^p(\Omega)$  for  $p \in [2, \frac{2n}{n-2k})$ .*

*Proof* Extend  $f \in H^k(\Omega)$  to  $E[f] \in H^k(\mathbb{R}^n)$ , apply the above result, then restrict back to  $\Omega$ ; clearly  $\|f\|_{L^p(\Omega)} \leq \|E[f]\|_{L^p(\mathbb{R}^n)}$ . □

### 2.9.3 Scaling and inequalities

We have (almost) shown that if  $f \in H^k(\mathbb{R}^n)$  then  $f \in L^p(\mathbb{R}^n)$  with  $p = 2n/(n-2k)$ . Why is this what we should expect?

Suppose that we do have an inequality

$$\|f\|_{L^p} \leq C\|f\|_{H^k}.$$

What happens if we rescale the function  $f(x)$  and consider instead  $f_\lambda(x) = f(\lambda x)$ . This should satisfy the same inequality

$$\|f_\lambda\|_{L^p} \leq C\|f_\lambda\|_{H^k}.$$

Now,

$$\|f_\lambda\|_{L^p}^p = \int_{\mathbb{R}^n} |f(\lambda x)|^p dx = \lambda^{-n} \int_{\mathbb{R}^n} |f(y)|^p dy,$$

and so

$$\|f_\lambda\|_{L^p} = \lambda^{-n/p} \|f\|_{L^p}.$$

Each derivative also brings out a factor of  $\lambda$ , and so

$$\|\partial^\alpha f\|_{L^2} = \lambda^{|\alpha|-n/2} \|f\|_{L^2}.$$

Putting  $|\alpha| = k$ , we would expect

$$k - \frac{n}{2} = -\frac{n}{p},$$

which gives  $p = 2n/(n - 2k)$ .

### 2.9.4 Boundary values and the trace operator

If  $\Omega$  is a bounded  $C^k$  domain then we say that  $f \in H^k(\partial\Omega)$  if

$$(\Phi_j^{-1})^*(\psi_j f) \in H^k(\mathbb{R}^{n-1})$$

for every  $j$ , where  $\psi_j$  is a partition of unity as constructed in the proof of Theorem 2.27, and define

$$\|f\|_{H^k(\partial\Omega)}^2 = \sum_{j=1}^m \|(\Phi_j^{-1})^*(\psi_j f)\|_{H^k(\mathbb{R}^{n-1})}^2.$$

(One can show that any choice of  $\Phi_j$  and  $\psi_j$  will give equivalent norms.)

**Theorem 2.40** *Suppose that  $\Omega$  is a bounded  $C^1$  domain. Then there exists a bounded linear operator*

$$T : H^1(\Omega) \rightarrow L^2(\partial\Omega)$$

*such that for every  $f \in H^1(\Omega) \cap C^0(\bar{\Omega})$ ,  $Tu = u|_{\partial\Omega}$ .*

*Proof* We first prove the result for  $f \in H^1(\mathbb{R}_+^n)$  and show that there is a bounded linear operator from  $H^1(\mathbb{R}_+^n) \cap C^1(\bar{\mathbb{R}}_+^n)$  into  $L^2(\partial\mathbb{R}_+^n) = L^2(\mathbb{R}^{n-1})$ . For any  $f \in H^1(\mathbb{R}_+^n) \cap C^1(\bar{\mathbb{R}}_+^n)$  such that  $f(x) \rightarrow 0$  as  $x_n \rightarrow \infty$ , we have

$$\begin{aligned} \int_{\partial\mathbb{R}_+^n} |f(x', 0)|^2 dx' &= - \int_{\mathbb{R}_+^n} \partial_n(|f|^2) dx \\ &= - \int_{\mathbb{R}_+^n} 2f \partial_n f dx \\ &\leq \int |f|^2 + |\partial_n f|^2 dx \\ &\leq \|f\|_{H^1(\mathbb{R}_+^n)}^2. \end{aligned}$$

Now for  $f \in H^1(\Omega) \cap C^1(\bar{\Omega})$ , we use the partition of unity and the above result for  $\mathbb{R}_+^n$  so that  $\|f\|_{L^2(\partial\Omega)} \leq \|f\|_{H^1(\Omega)}$ . Define  $Tf = f|_{\partial\Omega}$ .

Given any  $f \in H^1(\Omega)$ , there is a sequence  $f_n \in C^\infty(\bar{\Omega})$  that converges to  $f$  in  $H^1(\Omega)$ ; in particular  $f_n \in H^1(\Omega) \cap C^1(\bar{\Omega})$ , so we can define  $Tf$  by taking limits. Since  $f_n \rightarrow f$  uniformly on  $\Omega$  (see examples) the result holds as stated.  $\square$

**Theorem 2.41**  $f \in H_0^1(\Omega)$  iff  $f \in H^1(\Omega)$  and  $Tf = 0$ .

### 3

---

## Elliptic PDEs

We now prove existence and uniqueness result for the weak formulation of the elliptic problem

$$Lu = f \quad u|_{\partial\Omega} = 0, \quad (3.1)$$

where

$$Lu = - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^n b_i(x) \frac{\partial u}{\partial x_i} + c(x)u. \quad (3.2)$$

For simplicity we will suppose that the coefficients  $(a_{ij}, b_i, \text{ and } c)$  are smooth functions of  $x$ .

First we want to recast this equation in a weak form. If we multiply the equation by a function  $\varphi \in C_c^\infty(\Omega)$  and integrate by parts we obtain

$$\sum_{i,j=1}^n (a_{ij} \partial_j u, \partial_i \varphi) + \sum_{i=1}^n (b_i \partial_i u, \varphi) + (cu, \varphi) = (f, \varphi).$$

Write

$$B(u, \varphi) = \sum_{i,j=1}^n (a_{ij} \partial_j u, \partial_i \varphi) + \sum_{i=1}^n (b_i \partial_i u, \varphi) + (cu, \varphi),$$

and note that for a fixed smooth  $u$ , the left-hand side defines a linear map

from  $C_c^\infty(\Omega)$  into  $\mathbb{R}$ , such that

$$\begin{aligned} |B(u, \varphi)| &\leq \sum_{i,j=1}^n \|a_{ij} \partial_j u\|_{L^2} \|\partial_i \varphi\|_{L^2} + \sum_{i=1}^n \|b_i \partial_i u\|_{L^2} \|\varphi\|_{L^2} + \|cu\|_{L^2} \|\varphi\|_{L^2} \\ &\leq \sum_{i,j=1}^n \|a_{ij}\|_{L^\infty} \|\partial_j u\|_{L^2} \|\varphi\|_{H^1} + \sum_{i=1}^n \|b_i\|_{L^\infty} \|\partial_i u\|_{L^2} \|\varphi\|_{H^1} + \|c\|_{L^\infty} \|u\|_{L^2} \|\varphi\|_{H^1} \\ &\leq C(a_{ij}, b_i, c) \|u\|_{H^1} \|\varphi\|_{H^1}. \end{aligned}$$

Note also that if  $f \in L^2$  (say) then

$$|(f, \varphi)| \leq \|f\|_{L^2} \|\varphi\|_{L^2} \leq \|f\|_{L^2} \|\varphi\|_{H^1}.$$

Thus if  $B(u, \varphi) = (f, \varphi)$  for every  $\varphi \in C_c^\infty(\Omega)$  then it also holds for every  $\varphi \in H_0^1(\Omega)$ , using the density of  $C_c^\infty(\Omega)$  in  $H_0^1(\Omega)$ .

So we have shown that if  $u$  is a smooth solution of (3.1), then in fact we must have

$$B(u, \varphi) = (f, \varphi) \quad \text{for all } \varphi \in H_0^1(\Omega), \quad (3.3)$$

which is the weak form of (3.1). One can show – see examples for a simpler case – that if  $u$  is smooth and (3.3) holds then  $u$  satisfies (3.1) in the classical sense. We can weaken (3.3) just a little further, by noting that the right-hand side in fact defines a linear functional on  $H_0^1(\Omega)$ , and choosing to consider instead the problem

$$B(u, \varphi) = F(\varphi) \quad \text{for all } \varphi \in H_0^1(\Omega)$$

for some  $F \in H^{-1}(\Omega)$ .

The problem is now in the right form to apply the Lax–Milgram Lemma. We have already shown that our  $B$  is bounded from  $H_0^1 \times H_0^1$  into  $\mathbb{R}$ ; all that is left is to show that  $B$  is coercive, i.e. that there exists a  $\beta > 0$  such that

$$B(u, u) \geq \beta \|u\|_{H^1}^2 \quad \text{for all } u \in H_0^1(\Omega).$$

This isn't quite true; but we will show that there is a constant  $\lambda > 0$  such that  $B(u, v) + \lambda(u, v)$  is coercive.

In order to do this we will require the assumption that  $L$  is uniformly elliptic: there exists a  $\theta > 0$  such that

$$\sum_{i,j=1}^n a_{ij}(x) \xi_i \xi_j \geq \theta |\xi|^2. \quad (3.4)$$

So now

$$|B(u, u)| \geq \theta \sum_{j=1}^n \|\partial_j u\|_{L^2}^2 - \left( \max_i \|b_i\|_{L^\infty} \right) \|\nabla u\|_{L^2} \|u\|_{L^2} - \|c\|_{L^\infty} \|u\|_{L^2}^2.$$

We use “Young’s inequality with  $\epsilon$ ”,

$$ab \leq \frac{\epsilon}{2} a^2 + \frac{1}{2\epsilon} b^2$$

on the second term,

$$b \|\nabla u\|_{L^2} \|u\|_{L^2} \leq \frac{\theta}{2} \|\nabla u\|_{L^2}^2 + \frac{1}{2\theta} b^2 \|u\|_{L^2}^2,$$

so that

$$|B(u, u)| \geq \frac{\theta}{2} \|\nabla u\|_{L^2}^2 - \lambda' \|u\|_{L^2}^2, \quad (3.5)$$

where

$$\lambda' = \frac{1}{2\theta} \left( \max_i \|b_i\|_{L^\infty} \right)^2 + \|c\|_{L^\infty}.$$

So finally

$$|B(u, u)| \geq \frac{\theta}{2} \|u\|_{H^1}^2 - \lambda_0 \|u\|_{L^2}^2,$$

where  $\lambda_0 = \lambda' + (\theta/2)$ . (We could also use Poincaré’s inequality to get a better constant in front of the  $\|u\|_{L^2}^2$  term.)

We therefore obtain the following theorem:

**Theorem 3.1** *Let  $L$  be a uniformly elliptic operator, and  $B$  the associated bilinear form. Then there exists a constant  $\lambda_0 > 0$  such that for any  $\lambda \geq \lambda_0$  the equation*

$$B(u, \varphi) + \lambda(u, \varphi) = (f, \varphi) \quad \text{for all } \varphi \in H_0^1(\Omega)$$

has a unique solution  $u \in H_0^1(\Omega)$ , and

$$\|u\|_{H^1} \leq c \|f\|_{H^{-1}}.$$

*Proof* The Lax–Milgram Lemma can be applied to the bilinear form  $B'(u, v) = B(u, v) + \lambda(u, v)$ .  $\square$

Note that if  $b_i = c = 0$  then we can take  $\lambda_0 = 0$  in the previous theorem (and we would only need to use the Riesz representation theorem rather than the full Lax–Milgram Lemma).

But we should be able to do better if  $f$  is nicer...

### 3.1 Elliptic regularity

Suppose that  $f \in L^2(\Omega)$  and consider again Poisson’s equation  $-\Delta u = f$ . Suppose that  $u$  is smooth; then we can take the inner product of the equation with  $\partial_k^2 u$  and write

$$\begin{aligned} \int_{\Omega} (\Delta u)(\partial_k^2 u) \, dx &= \sum_{i=1}^n \int_{\Omega} (\partial_i^2 u)(\partial_k^2 u) \, dx \\ &= - \sum_{i=1}^n \int_{\Omega} (\partial_i u)(\partial_i \partial_k^2 u) \, dx + \sum_{i=1}^n \int_{\partial\Omega} (\partial_i u)(\partial_k^2 u) n_i \, dS \\ &= \sum_{i=1}^n \int_{\Omega} (\partial_i \partial_k u)^2 \, dx \\ &\quad + \sum_{i=1}^n \int_{\partial\Omega} (\partial_i u)(\partial_k^2 u) n_i - (\partial_i u)(\partial_i \partial_k u) n_k \, dS, \end{aligned}$$

Summing over  $k$  this yields

$$\sum_{i,k=1}^n \|\partial_i \partial_k u\|_{L^2}^2 = \|\Delta u\|_{L^2}^2 + \text{boundary terms},$$

i.e. the second derivatives are controlled by the Laplacian. Since  $-\Delta u = f$  this yields

$$\sum_{|\alpha|=2} \|\partial^\alpha u\|_{L^2}^2 = \|f\|_{L^2}^2 + \text{boundary terms}.$$

#### 3.1.1 Interior regularity

It is relatively easy to show that if  $\Omega' \subset\subset \Omega$  then the  $H^2(\Omega')$  of  $u$  is bounded by the  $L^2(\Omega)$  norm of  $f$  (‘interior regularity’). We give a less-than rigorous proof, which nevertheless contains the main idea.

**Theorem 3.2** *If  $f \in L^2(\Omega)$ , and  $u \in H_0^1(\Omega)$  satisfies  $Lu = f$ , where  $L$  is a uniformly elliptic operator as in (3.2) (plus (3.4)), then for every  $\Omega' \subset\subset \Omega$  there exists a constant  $c_{\Omega'}$*

$$\|u\|_{H^2(\Omega')} \leq c_{\Omega'} (\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)}). \quad (3.6)$$

Of course, if we know where  $u$  comes from (for example, Theorem 3.1), then we already have  $\|u\|_{H^1(\Omega)} \leq c\|f\|_{L^2(\Omega)}$ , and (3.6) becomes

$$\|u\|_{H^2(\Omega')} \leq C_{\Omega'} \|f\|_{L^2(\Omega)}. \quad (3.7)$$

Another remark is in order; if one looks carefully at the proof, we do not in fact need the  $L^2$  norm of  $f$  on the whole of  $\Omega$ , but only on a compact subset of  $\Omega$ , say  $\hat{\Omega}$  that contains  $\Omega'$ . This means that if  $f \in C^0(\Omega)$ , then  $f \in L^2(\hat{\Omega})$ , and hence  $u \in H^2(\Omega')$ . This observation will be useful below in Corollary 3.4.

*Proof* Here is a non-rigorous proof for the simple case of Poisson's equation. Take  $\Omega' \subset\subset \Omega$ , and choose a cut-off function  $\zeta \in C_c^\infty(\Omega)$  such that  $0 \leq \zeta \leq 1$  and  $\zeta \equiv 1$  on  $\Omega'$ . Now, assuming that  $u$  is smooth,

$$\begin{aligned} \int_{\Omega} \zeta^2 |\Delta u|^2 dx &= \sum_{i,j} \int_{\Omega} (\partial_i^2 u)(\partial_j^2 u) \zeta^2 dx \\ &= - \sum_{i,j} \int_{\Omega} (\partial_i u)(\partial_i \partial_j^2 u) \zeta^2 dx - 2 \sum_{i,j} \int_{\Omega} (\partial_i u)(\partial_j^2 u) \zeta (\partial_i \zeta) dx \\ &= \sum_{i,j} \int_{\Omega} (\partial_i \partial_j u)^2 \zeta^2 dx \\ &\quad + 2 \sum_{i,j} \int_{\Omega} (\partial_i u)(\partial_i \partial_j u) \zeta (\partial_j \zeta) - (\partial_i u)(\partial_j^2 u) \zeta (\partial_i \zeta) dx. \end{aligned}$$

(Note that this is the same as multiplying by the  $n$  'test functions'  $\zeta^2 \partial_j^2 u$  and then summing over  $j$ .) The left-hand side is bounded by

$$\int_{\Omega} \zeta^2 |\Delta u|^2 dx \leq \int_{\Omega} |\Delta u|^2 dx = \int_{\Omega} |f|^2 dx,$$

so we just have to estimate the messy looking terms (with derivatives of  $\zeta$ ). Since  $\zeta$  is fixed and smooth, we have  $\|\partial_j \zeta\|_{L^\infty}^2 \leq c$ , so

$$\begin{aligned} \left| \int_{\Omega} (\partial_i u) (\partial_i \partial_j u) \zeta (\partial_j \zeta) \, dx \right| &\leq \|\partial_j \zeta\|_{L^\infty} \int_{\Omega} |\partial_i u| |\partial_i \partial_j u| |\zeta| \, dx \\ &\leq 2n \|\partial_j \zeta\|_{L^\infty}^2 \int_{\Omega} |\partial_i u|^2 \, dx + \frac{1}{8n} \int_{\Omega} (\partial_i \partial_j u)^2 \zeta^2 \, dx; \end{aligned}$$

we've used  $ab \leq 2na^2 + b^2/8n$  on the integrand here.

So we have

$$\begin{aligned} \sum_{i,j} \int_{\Omega} (\partial_i \partial_j u)^2 \zeta^2 \, dx &\leq \|f\|_{L^2(\Omega)}^2 + \sum_{i,j} 4cn \|\partial_i u\|_{L^2(\Omega)}^2 + \frac{1}{4n} \int_{\Omega} (\partial_i \partial_j u)^2 \zeta^2 \\ &\quad + \sum_{i,j} 4cn \|\partial_i u\|_{L^2(\Omega)}^2 + \frac{1}{4n} \int_{\Omega} (\partial_j^2 u) \zeta^2 \, dx \\ &\leq \|f\|_{L^2(\Omega)}^2 + 8cn^2 \|u\|_{H^1(\Omega)}^2 + \frac{1}{2} \int_{\Omega} (\partial_i \partial_j u)^2 \zeta^2 \, dx, \end{aligned}$$

from which it follows that

$$\begin{aligned} \sum_{|\alpha|=2} \|\partial_\alpha u\|_{L^2(\Omega')}^2 &= \sum_{i,j=1}^n \|\partial_i \partial_j u\|_{L^2(\Omega')}^2 \\ &\leq \sum_{i,j} \int_{\Omega} (\partial_i \partial_j u)^2 \zeta^2 \, dx \\ &\leq 2\|f\|_{L^2(\Omega)}^2 + 16cn^2 \|u\|_{H^1(\Omega)}^2, \end{aligned}$$

and the result follows since

$$\|u\|_{H^2(\Omega')}^2 = \sum_{|\alpha|=2} \|\partial_\alpha u\|_{L^2(\Omega')}^2 + \|u\|_{H^1(\Omega')}^2.$$

□

The lack of rigour is due to the fact that our solution  $u$  does not satisfy the equation  $-\Delta u = f$ , but the weak form  $(\nabla u, \nabla v) = (f, v)$  for every  $v \in H_0^1(\Omega)$ , and that as far as we know  $u$  is only in  $H_0^1(\Omega)$  and isn't smooth. How do we get round this? We want to do something like choosing  $v = \zeta^2 \Delta u$  as a test function, but since we only have  $u \in H_0^1$ , this isn't smooth enough (we'd have  $v \in H^{-1}$ , in fact).

In order to deal with this we use difference quotients; given a function  $u$  we define

$$D_i^h u(x) = \frac{u(x + he_i) - u(x)}{h},$$

where  $e_i$  is a unit vector in the  $i^{\text{th}}$  direction. Then one can show that if  $u \in H^1(\Omega)$ , for any  $\Omega' \subset\subset \Omega$

$$\|D_i^h u\|_{L^2(\Omega')} \leq \|\partial_i u\|_{L^2(\Omega)},$$

i.e. ‘differences look like derivatives’, and more importantly that if  $\Omega' \subset\subset \Omega$  and for every  $h < \text{dist}(\Omega', \partial\Omega)$  we have

$$\|D_i^h u\|_{L^2(\Omega')} \leq C,$$

then  $\partial_i u \in L^2(\Omega')$  and

$$\|\partial_i u\|_{L^2(\Omega')} \leq C.$$

In other words, we can recover weak derivatives by the same sort of limiting process that we can define classical derivatives.

In the rigorous version of the proof one takes the  $n$  test functions  $v = D_k^{-h}(\zeta^2 D_k^h v)$  with  $k = 1, \dots, n$ ; note that these *are* elements of  $H_0^1(\Omega)$  if  $h$  is small enough. Essentially one then repeats the estimates in the above theorem in this setting, showing that the estimates are uniform when  $h$  is sufficiently small.

If  $f$  is smoother, then  $u$  is smoother:

**Theorem 3.3** *Let  $u \in H_0^1(\Omega)$  satisfy  $Lu = f$ , as in the above theorem. If  $f \in H^s(\Omega)$  then  $u \in H_{\text{loc}}^{s+2}(\Omega)$ ; for any  $\Omega' \subset\subset \Omega$  there exists a constant  $c_{\Omega', s}$  such that*

$$\|u\|_{H^{s+2}(\Omega')} \leq c \|f\|_{H^s(\Omega)}.$$

*Proof* Again, we dispense with rigour: suppose that  $-\Delta u = f$ . Then the derivative  $\partial^\alpha u$  satisfies

$$-\Delta(\partial^\alpha u) = \partial^\alpha f.$$

If  $f \in H^s(\Omega)$ ,  $\partial^\alpha f \in L^2(\Omega)$  for any  $\alpha$  with  $|\alpha| = s$ ; so  $\partial^\alpha u \in H^2(\Omega')$  by the preceding interior regularity result. In the rigorous version, one chooses test functions  $v = (-1)^{|\alpha|} \partial^\alpha w$  with  $w \in C_c^\infty(\Omega)$  and with  $|\alpha| = s + 1$ , see examples.  $\square$

**Corollary 3.4** *If  $f \in C^\infty(\Omega)$  and  $u \in H_0^1(\Omega)$  is a weak solution of  $Lu = f$  then  $u \in C^\infty(\Omega)$ .*

*Proof* Take  $\Omega' \subset\subset \Omega'' \subset\subset \Omega$ . If  $f \in C^\infty(\Omega)$  then  $f \in C^\infty(\bar{\Omega}'')$ . It follows that  $f \in H^k(\Omega'')$  for every  $k \in \mathbb{N}$ . It then follows that  $u \in H^{k+2}(\Omega')$  for every  $k \in \mathbb{N}$ , and hence in  $C^r(\bar{\Omega}')$  for every  $r$ . Since this holds for every  $\Omega' \subset\subset \Omega$ ,  $u \in C^\infty(\Omega)$ .  $\square$

**Corollary 3.5** *If  $u \in H_0^1(\Omega)$  satisfies the weak form of*

$$Lu = \lambda u$$

*then  $u \in C^\infty(\Omega)$ , and hence is a classical solution of  $Lu = \lambda u$ .*

*Proof* Suppose that  $u \in H^r(\Omega')$  for any  $\Omega' \subset\subset \Omega$ . Now pick a particular  $\Omega' \subset\subset \Omega$ , and find  $\Omega''$  such that  $\Omega' \subset\subset \Omega'' \subset\subset \Omega$ . We know that  $u \in H^r(\Omega'')$ , and then  $u \in H^{r+2}(\Omega')$ . Induction finishes the proof.  $\square$

### 3.1.2 Boundary regularity

To show that if  $f \in L^2(\Omega)$  then in fact  $u \in H^2(\Omega)$  requires more work; but the idea is clear if one considers the simple case of  $-\Delta u = f$  in the upper half plane, which we treat now.

Suppose that  $u$  is smooth and has compact support in  $\overline{\mathbb{R}_+^n}$ , contained in  $D^+ = B(0, R) \cap \overline{\mathbb{R}_+^n}$ , and that  $u$  satisfies

$$(\nabla u, \nabla v) = (g, v) \quad \text{for all } v \in H_0^1(D^+),$$

where  $D^+$  is  $B(0, R) \cap \mathbb{R}_+^n$ . Then

$$\|u\|_{H^2(\mathbb{R}_+^n)} \leq c \|g\|_{\mathbb{R}_+^n}.$$

Take  $v = \partial_k^2 u$  as a test function. If  $k \neq n$  then note that  $v$  is still zero on  $x_n = 0$ , so is still an allowable test function. Then

$$\begin{aligned} \int_{\mathbb{R}_+^n} g \partial_k^2 u &= \sum_i \int_{\mathbb{R}_+^n} (\partial_i u) (\partial_i \partial_k^2 u) \, dx \\ &= - \sum_i \int_{\mathbb{R}_+^n} (\partial_i \partial_k u) (\partial_i \partial_k u) \, dx, \end{aligned}$$

where we have integrated by parts with respect to  $x_k$ . Since  $k \neq n$ , we do not include any terms from the boundary  $x_n = 0$ ; the boundary terms that we would pick up vanish.

Now use Cauchy-Schwarz:

$$\begin{aligned} \sum_i \|\partial_k \partial_i u\|_{L^2(\mathbb{R}_+^n)}^2 &\leq \|g\|_{L^2(\mathbb{R}_+^n)} \|\partial_k^2 u\|_{L^2(\mathbb{R}_+^n)} \\ &\leq \|g\|_{L^2(\mathbb{R}_+^n)} \left( \sum_i \|\partial_k \partial_i u\|_{L^2(\mathbb{R}_+^n)} \right), \end{aligned}$$

whence

$$\sum_i \|\partial_k \partial_i u\|_{L^2(\mathbb{R}_+^n)}^2 \leq \|g\|_{L^2(\mathbb{R}_+^n)}^2.$$

This shows that every second partial derivative apart from  $\partial_n^2 u$  is in  $L^2(\mathbb{R}_+^n)$ . But the equation  $-\Delta u = f$  shows that

$$\partial_n^2 u = f - \sum_{j=1}^{n-1} \partial_j^2 u.$$

Since each term on the right-hand side is in  $L^2(\mathbb{R}_+^n)$ , so is the left-hand side.

To make this work rigorously, one uses a cutoff function, a partition of unity, and a straightening of the boundary. Note that when one uses the coordinate transformation to straighten the boundary, one necessarily introduces other terms in the equation (even if we start with just the Laplacian, we end up having to consider a more general elliptic equation). We state the theorem, including the higher regularity result.

**Theorem 3.6** *Suppose that  $\Omega$  is a bounded domain with a  $C^{r+2}$  boundary. Then if  $u \in H_0^1(\Omega)$  is a weak solution of  $Lu = f$  and  $f \in H^r(\Omega)$  then in fact  $u \in H^{r+2}(\Omega)$  and*

$$\|u\|_{H^2(\Omega)} \leq c(\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)}).$$

**Corollary 3.7** *Suppose that for some  $\lambda > 0$ , there exists a solution  $u \in H_0^1(\Omega)$  of the eigenvalue problem*

$$Lu = \lambda u.$$

*Then  $u \in C^\infty(\bar{\Omega})$ .*

---

## Eigenvalues, eigenfunctions, and bases in Hilbert spaces

### 4.1 Eigenvalues and the spectrum

Let  $H$  be a Hilbert space and  $T \in B(H, H)$ , then the *point spectrum* of  $T$  consists of the set of all eigenvalues,

$$\sigma_p(T) = \{\lambda \in \mathbb{C} : Tx = \lambda x \text{ for some non-zero } x \in H\}.$$

If  $A$  is a linear operator on a finite-dimensional space  $V$  then  $\lambda \in \mathbb{C}$  is an eigenvalue of  $A$  if

$$Ax = \lambda x \text{ for some non-zero } x \in V.$$

In this case  $\lambda$  is an eigenvalue if and only if  $A - \lambda I$  is not invertible (recall that you can find the eigenvalues of an  $n \times n$  matrix by solving  $\det(A - \lambda I) = 0$ ). However, this is no longer true in infinite-dimensional spaces, and the *spectrum* of  $A$  is potentially larger than the set of eigenvalues.

**Definition 4.1** *The resolvent set of  $T$ ,  $R(T)$ , is*

$$R(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ has a bounded inverse defined on all of } H\}.$$

*The spectrum  $\sigma(T)$  of  $T \in B(H, H)$  is the complement of  $R(T)$ ,*

$$\sigma(T) = \mathbb{C} \setminus R(T),$$

*i.e. the spectrum of  $T$  is the set of all complex  $\lambda$  for which  $T - \lambda I$  does not have a bounded inverse defined on all of  $H$ .*

A bounded linear operator  $T \in \mathcal{L}(H, H)$  is symmetric iff

$$(Tu, v) = (u, Tv) \quad \text{for all } u, v \in H.$$

For compact symmetric operators, the spectrum of  $T$  consists only of the eigenvalues of  $T$  (and perhaps zero). We will not prove this here, and content ourselves with investigating the eigenvalue problem.

## 4.2 The spectrum of a compact symmetric operator

**Theorem 4.2** *Let  $T \in \mathcal{L}(H, H)$  be symmetric. Then all the eigenvalues of  $T$  are real and the eigenvectors corresponding to distinct eigenvalues are orthogonal.*

*Proof* Suppose that  $Tx = \lambda x$  with  $x \neq 0$ . Then

$$\lambda \|x\|^2 = (\lambda x, x) = (Tx, x) = (x, T^*x) = (x, Tx) = (x, \lambda x) = \bar{\lambda} \|x\|^2,$$

i.e.  $\lambda = \bar{\lambda}$ .

Now if  $\lambda$  and  $\mu$  are distinct eigenvalues with  $Tx = \lambda x$  and  $Ty = \mu y$  then

$$0 = (Tx, y) - (x, Ty) = (\lambda x, y) - (x, \mu y) = (\lambda - \mu)(x, y),$$

and so  $(x, y) = 0$ . □

It is not immediately obvious that the following result has anything to do with eigenvalues.

**Theorem 4.3** *Let  $H$  be a Hilbert space and  $T \in \mathcal{L}(H, H)$  a symmetric operator. Then*

- (a)  $(Tx, x)$  is real for all  $x \in H$  and
- (b)  $\|T\| = \sup\{|(Tx, x)| : x \in H, \|x\| = 1\}$ .

*Proof* For (a) we have

$$(Tx, x) = (x, Tx) = \overline{(Tx, x)},$$

and so  $(Tx, x)$  is real. Now let  $M = \sup\{|(Tx, x)| : x \in H, \|x\| = 1\}$ . Clearly

$$|(Tx, x)| \leq \|Tx\| \|x\| \leq \|T\| \|x\|^2 = \|T\|$$

when  $\|x\| = 1$ , and so  $M \leq \|T\|$ .

For any  $u, v \in H$  we have

$$\begin{aligned} 4(Tu, v) &= (T(u+v), u+v) - (T(u-v), u-v) \\ &\leq M(\|u+v\|^2 + \|u-v\|^2) \\ &\leq 2M(\|u\|^2 + \|v\|^2) \end{aligned}$$

using the parallelogram law. If  $Tu \neq 0$  choose

$$v = \frac{\|u\|}{\|Tu\|} Tu$$

to obtain, since  $\|v\| = \|u\|$ , that

$$4\|u\|\|Tu\| \leq 4M\|u\|^2,$$

i.e.  $\|Tu\| \leq M\|u\|$ . This also holds if  $Tu = 0$ . It follows that  $\|T\| \leq M$  and therefore that  $\|T\| = M$ .  $\square$

Clearly  $|\lambda| \leq \|T\|_{\text{op}}$  for any  $\lambda \in \sigma_p(T)$ , since if  $Tx = \lambda x$  then

$$|\lambda|\|x\| = |\lambda x| = \|Tx\| \leq \|T\|_{\text{op}}\|x\|.$$

We now consider eigenvalues of compact self-adjoint linear operators on a Hilbert space. It is convenient to restrict attention to Hilbert spaces over  $\mathbb{C}$ , but this is no restriction, since we can always consider the ‘complexification’ of a real Hilbert space<sup>1</sup>.

<sup>1</sup> Let  $H$  be a Hilbert space over  $\mathbb{R}$ , and define its complexification  $H_{\mathbb{C}}$  as the vector space

$$H_{\mathbb{C}} = \{x + iy : x, y \in H\},$$

equipped with operations  $+$  and  $*$  defined via

$$(x + iy) + (w + iz) = (x + w) + i(y + z), \quad x, y, w, z \in V$$

and

$$(a + ib) * (x + iy) = (ax - by) + i(bx + ay) \quad a, b \in \mathbb{R}, \quad x, y \in V.$$

Then equipped with the inner product

$$(x + iy, w + iz)_{H_{\mathbb{C}}} = (x, w) + i(y, w) - i(x, z) + (y, z)$$

$H_{\mathbb{C}}$  is a Hilbert space. Just as we can complexify a Hilbert space  $H$  to give  $H_{\mathbb{C}}$ , we can complexify a linear operator  $T$  that acts on  $H$  to a linear operator  $T_{\mathbb{C}}$  that acts on  $H_{\mathbb{C}}$ : given  $T \in B(H, H)$ , extend  $T$  to a linear operator  $\tilde{T} : H_{\mathbb{C}} \rightarrow H_{\mathbb{C}}$  via the definition

$$\tilde{T}(x + iy) = Tx + iTy \quad x, y \in H.$$

Then  $\tilde{T} \in B(H_{\mathbb{C}}, H_{\mathbb{C}})$ , any eigenvalue of  $T$  is an eigenvalue of  $\tilde{T}$ , and any real eigenvalue of  $\tilde{T}$  is an eigenvalue of  $T$ . If  $T$  is symmetric then  $\tilde{T}$  is symmetric.

Let's think again about the weak formulation of the elliptic problem; given  $f \in L^2$ , find  $u \in H_0^1(\Omega)$  such that

$$B(u, v) = (f, v) \quad \text{for all } v \in H_0^1(\Omega).$$

We could define a 'solution mapping' for this problem: if  $f \in L^2(\Omega)$  then  $Tf = u$ , where  $u$  is the solution.

We can observe that

- (i)  $T$  is linear; and
- (ii)  $T$  maps  $L^2(\Omega)$  into  $H_0^1(\Omega)$  continuously, i.e.

$$\|Tf\|_{H_0^1(\Omega)} \leq c\|f\|_{L^2(\Omega)}.$$

(In fact our elliptic regularity tells us more, than  $T$  maps  $L^2$  into  $H^2$  continuously.)

Since  $H_0^1(\Omega) \subset L^2(\Omega)$ ,  $T$  is clearly a linear map from  $L^2$  into itself, with the additional property that a bounded set in  $L^2$  is mapped into a bounded set in  $H_0^1$ . Since  $H^1$  is compactly embedded in  $L^2$ , this means that a bounded set in  $L^2$  becomes a (pre)compact set in  $L^2$  (precompact means that its closure is compact).  $T$  is therefore an example of a compact map:

**Definition 4.4** *A map  $T : X \rightarrow Y$  is compact if whenever  $U$  is a bounded subset of  $X$ , the closure of  $T(U)$  is a compact subset of  $Y$ . Equivalently, if  $\{x_n\}$  is a bounded sequence in  $X$ , then  $\{Tx_n\}$  has a subsequence that converges in  $Y$ .*

Note that a compact operator must be bounded, since otherwise there exists a sequence in  $H$  with  $\|x_n\| = 1$  but  $\|Tx_n\| \rightarrow \infty$ , and clearly  $\{Tx_n\}$  cannot have a convergent subsequence.

If  $T \in \mathcal{L}(H, K)$  has finite-dimensional range then  $T$  is compact, since any bounded sequence in a finite-dimensional space has a convergent subsequence.

We now show that any compact symmetric operator has at least one eigenvalue.

**Theorem 4.5** *Let  $H$  be a Hilbert space and  $T \in \mathcal{L}(H, H)$  a compact symmetric operator. Then at least one of  $\pm\|T\|_{\text{op}}$  is an eigenvalue of  $T$ .*

*Proof* We assume that  $T \neq 0$ , otherwise the result is trivial. From Theorem 4.3,

$$\|T\|_{\text{op}} = \sup_{\|x\|=1} |(Tx, x)|.$$

Thus there exists a sequence  $x_n$ , of unit vectors, such that

$$(Tx_n, x_n) \rightarrow \pm\|T\|_{\text{op}} = \alpha.$$

Since  $T$  is compact there is a subsequence  $x_{n_j}$  such that  $Tx_{n_j}$  is convergent to some  $y$ . Relabel  $x_{n_j}$  as  $x_n$  again.

Now consider

$$\begin{aligned} \|Tx_n - \alpha x_n\|^2 &= \|Tx_n\|^2 + \alpha^2 - 2\alpha(Tx_n, x_n) \\ &\leq 2\alpha^2 - 2\alpha(Tx_n, x_n); \end{aligned}$$

by the choice of  $x_n$ , the right-hand side tends to zero as  $n \rightarrow \infty$ . It follows, since  $Tx_n \rightarrow y$ , that

$$\alpha x_n \rightarrow y,$$

and since  $\alpha \neq 0$  is fixed we must have  $x_n \rightarrow x$  for some  $x \in H$ . Therefore  $Tx_n \rightarrow Tx = \alpha x$ . It follows that

$$Tx = \alpha x$$

and clearly  $x \neq 0$ , since  $\|y\| = |\alpha|\|x\| = \|T\|_{\text{op}} \neq 0$ .  $\square$

Note that since any eigenvalue must satisfy  $|\lambda| \leq \|T\|_{\text{op}}$ , it follows that for compact symmetric operators

$$\|T\|_{\text{op}} = \sup\{|\lambda| : \lambda \in \sigma_p(T)\}.$$

**Proposition 4.6** *Let  $T$  be a compact self-adjoint operator on a Hilbert space  $H$ . Then  $\sigma_p(T)$  is either finite or consists of a countable sequence tending to zero.*

*Proof* Suppose that  $T$  has infinitely many eigenvalues that do not form a sequence tending to zero. Then for some  $\epsilon > 0$  there exists a sequence of distinct eigenvalues with  $|\lambda_n| > \epsilon$ . Let  $x_n$  be a corresponding sequence of eigenvectors with  $\|x_n\| = 1$ ; then

$$\|Tx_n - Tx_m\|^2 = (Tx_n - Tx_m, Tx_n - Tx_m) = |\lambda_n|^2 + |\lambda_m|^2 \geq 2\epsilon^2$$

since  $(x_n, x_m) = 0$ . It follows that  $\{Tx_n\}$  can have no convergent subsequence, which contradicts the compactness of  $T$ .  $\square$

### 4.3 Bases in Hilbert spaces

In an infinite-dimensional separable Hilbert space the best that we can hope for in terms of a basis is to find a countable set  $\{e_j\}_{j=1}^{\infty}$ , in terms of which to expand any  $x \in H$  as potentially infinite series<sup>1</sup>,

$$x = \sum_{j=1}^{\infty} \alpha_j e_j. \quad (4.1)$$

A set  $\{e_j\}_{j=1}^{\infty}$  is a *basis* for  $H$  if every  $x \in H$  can be written uniquely in the form (4.1) for some  $\alpha_j \in \mathbb{K}$ . If in addition  $\{e_j\}_{j=1}^{\infty}$  is an orthonormal set (i.e.  $(e_i, e_j) = \delta_{ij}$ ) then we refer to it as an *orthonormal basis*. We concentrate on orthonormal bases.

**Lemma 4.7** *Let  $H$  be a Hilbert space and  $\{e_j\}_{j=1}^{\infty}$  an orthonormal sequence in  $H$ . Then for any  $x \in H$*

$$\sum_{j=1}^{\infty} |(x, e_j)|^2 \leq \|x\|^2$$

(‘Bessel’s inequality’); consequently for any  $x \in H$  the series

$$\sum_{n=1}^{\infty} (x, e_n) e_n$$

converges to some  $y \in H$  with  $\|y\|^2 = \sum_{j=1}^{\infty} |(x, e_j)|^2$ .

For  $\{e_n\}$  to be a basis we need (of course)  $y = x$ .

*Proof* Let us denote by  $x_k$  the partial sum

$$x_k = \sum_{j=1}^k (x, e_j) e_j \quad \text{with} \quad \|x_k\|^2 = \sum_{j=1}^k |(x, e_j)|^2.$$

Therefore

$$\begin{aligned} \|x - x_k\|^2 &= (x - x_k, x - x_k) \\ &= \|x\|^2 - (x_k, x) - (x, x_k) + \|x_k\|^2 \\ &= \|x\|^2 - \sum_{j=1}^k (x, e_j)(e_j, x) - \sum_{j=1}^k \overline{(x, e_j)}(x, e_j) + \|x_k\|^2 \\ &= \|x\|^2 - \|x_k\|^2. \end{aligned}$$

<sup>1</sup> equality here means that the partial sums converge to  $x$  in the norm of  $H$

It follows that for every  $k$

$$\sum_{j=1}^k |(x, e_j)|^2 = \|x_k\|^2 \leq \|x\|^2 - \|x - x_k\|^2 \leq \|x\|^2.$$

Thus if  $m > n$

$$\|x_m - x_n\|^2 = \left\| \sum_{j=n+1}^m (x, e_j)e_j \right\|^2 = \sum_{j=n+1}^m |(x, e_j)|^2,$$

which shows that  $\{x_m\}$  is a Cauchy sequence converging to sum  $y$ ; the expression for the norm follows since the norm of the limit is the limit of the norms.  $\square$

We now show give criteria for  $\{e_n\}$  to form a basis for  $H$ .

**Proposition 4.8** *Let  $E = \{e_j\}_{j=1}^{\infty}$  be an orthonormal set in a Hilbert space  $H$ . Then the following are equivalent to the statement that  $E$  is an orthonormal basis for  $H$ :*

- (a)  $x = \sum_{n=1}^{\infty} (x, e_n)e_n$  for all  $x \in H$ ;
- (b)  $\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2$  for all  $x \in H$ ; and
- (c)  $(x, e_n) = 0$  for all  $n$  implies that  $x = 0$ .
- (d) the linear span of  $\{e_n\}_{n=1}^{\infty}$  (the set of all finite linear combinations) is dense in  $H$ , i.e. for any  $\epsilon > 0$  there exist  $N$  and  $\alpha_j, j = 1, \dots, N$ , such that

$$\|x - \sum_{j=1}^N \alpha_j e_j\| < \epsilon.$$

*Proof* If  $E$  is an orthonormal basis for  $H$  then we can write

$$x = \sum_{j=1}^{\infty} \alpha_j e_j, \quad \text{i.e.} \quad x = \lim_{j \rightarrow \infty} \sum_{j=1}^n \alpha_j e_j.$$

Clearly if  $k \leq n$  we have

$$\left( \sum_{j=1}^n \alpha_j e_j, e_k \right) = \alpha_k,$$

and using the properties of the inner product of limits we obtain  $\alpha_k = (x, e_k)$  and hence (a) holds. The same argument shows that if we assume (a) then this expansion is unique and so  $E$  is a basis.

- (a)  $\Rightarrow$  (b) is immediate from Lemma 4.7.  
 (b)  $\Rightarrow$  (c) is immediate since  $\|x\| = 0$  implies that  $x = 0$ .  
 (c)  $\Rightarrow$  (a) Take  $x \in H$  and let

$$y = x - \sum_{j=1}^{\infty} (x, e_j) e_j.$$

For each  $m \in \mathbb{N}$  we have, using the continuity of the inner product,

$$\begin{aligned} (y, e_m) &= (x, e_m) - \lim_{n \rightarrow \infty} \left( \sum_{j=1}^n (x, e_j) e_j, e_m \right) \\ &= 0 \end{aligned}$$

since eventually  $n \geq m$ . It follows from (c) that  $y = 0$ .

Finally, it is clear that (a)  $\Rightarrow$  (d). We show that (d)  $\Rightarrow$  (c): suppose that  $(x, e_j) = 0$  for every  $j$ , and choose  $x_n$  contained in the linear span of the  $\{e_j\}$  such that  $x_n \rightarrow x$ . Then

$$\|x\|^2 = (x, x) = (x, \lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} (x, x_n) = 0,$$

since  $x_n$  is a finite linear combination of the  $\{e_j\}$ .  $\square$

**Lemma 4.9 (Gram–Schmidt orthonormalisation)** *If  $\{e_j\}_{j=1}^{\infty}$  is a linearly independent set, then there exists an orthonormal set  $\{\tilde{e}_j\}_{j=1}^{\infty}$  such that*

$$\text{span}(\tilde{e}_1, \dots, \tilde{e}_k) = \text{span}(e_1, \dots, e_k)$$

for every  $k \in \mathbb{N}$ .

*Proof* The proof is inductive. Suppose that we have already found  $\{\tilde{e}_1, \dots, \tilde{e}_n\}$ . Then set

$$\hat{e}_{n+1} = e_{n+1} - \sum_{j=1}^n (e_{n+1}, \tilde{e}_j) \tilde{e}_j \quad \text{then} \quad \tilde{e}_{n+1} = \frac{\hat{e}_{n+1}}{\|\hat{e}_{n+1}\|}.$$

It is easy to check that the new set  $\{\tilde{e}_1, \dots, \tilde{e}_{n+1}\}$  has the desired properties; and the induction starts with  $\tilde{e}_1 = e_1/\|e_1\|$ .  $\square$

**Proposition 4.10** *A Hilbert space  $H$  is separable iff it has a countable orthonormal basis.*

*Proof* Given a countable orthonormal basis, consider finite linear combinations with rational coefficients. This is countable and dense.

Given a countable dense subset, remove elements so that  $\{w_j\}_{j=1}^{\infty}$  is a linearly independent set whose closed linear span (elements that can be approximated by finite linear combinations) is all of  $H$ . Then apply the Gram–Schmidt process to produce an orthonormal set whose closed linear span is still all of  $H$ .  $\square$

**Theorem 4.11** *Any infinite-dimensional separable Hilbert space is isometric to  $\ell^2$  (i.e. there exists a linear isomorphism that preserves the norm).*

*Proof*  $H$  has a countable orthonormal basis  $\{e_j\}_{j=1}^{\infty}$ . Given  $x \in H$ ,  $x = \sum_{j=1}^{\infty} (x, e_j)e_j$ , and  $\|x\|^2 = \sum_{j=1}^{\infty} |(x, e_j)|^2$ . Set  $x_j = (x, e_j)$ , and consider the sequence  $i(x) = \underline{x} = (x_1, x_2, x_3, \dots)$ . Clearly  $\underline{x} \in \ell^2$  and  $\|\underline{x}\|_{\ell^2} = \|x\|$ . The map  $i$  is linear and invertible, with  $i^{-1}(\underline{x}) = \sum_{j=1}^{\infty} x_j e_j$ .  $\square$

#### 4.4 Eigenfunctions as basis elements; the Hilbert–Schmidt Theorem

We can use the Gram–Schmidt idea to show:

**Lemma 4.12** *If  $T \in \mathcal{L}(H, H)$  is compact and symmetric, then every non-zero eigenvalue has only a finite number of linearly independent eigenvectors.*

*Proof* Suppose that for some eigenvalue  $\lambda \neq 0$  there exists an infinite number of linearly independent eigenvectors  $\{w_j\}_{j=1}^{\infty}$ . Using the Gram–Schmidt process we can find a countably infinite collection of orthonormal eigenvectors  $\{e_j\}_{j=1}^{\infty}$  (any linear combination of the  $w_j$  is still an eigenvector with the same eigenvalue, since  $T(\sum_j \alpha_j w_j) = \sum_j \alpha_j T w_j = \lambda(\sum_j \alpha_j w_j)$ ). The proof now follows that of Proposition 4.6.  $\square$

In order to prove the main theorem on eigenvalues and eigenfunctions of compact symmetric operators we will require the following simple lemma.

**Lemma 4.13** *Let  $T \in \mathcal{L}(H, H)$  be symmetric and let  $S$  be a closed linear subspace of  $H$  such that  $TS \subseteq S$ . Then  $TS^{\perp} \subseteq S^{\perp}$ .*

*Proof* Let  $x \in S^\perp$  and  $y \in S$ . Then  $Ty \in S$  and so  $(Ty, x) = (y, Tx) = 0$  for all  $y \in S$ , i.e.  $Tx \in S^\perp$ .  $\square$

**Theorem 4.14** (*Hilbert–Schmidt Theorem*). *Let  $H$  be a Hilbert space and  $T \in B(H, H)$  be a compact self-adjoint operator. Then there exists a finite or countably infinite orthonormal sequence  $\{w_n\}$  of eigenvectors of  $T$  with corresponding non-zero real eigenvalues  $\{\lambda_n\}$  such that for all  $x \in H$*

$$Tx = \sum_j \lambda_j(x, w_j)w_j. \quad (4.2)$$

*Consequently there exists an orthonormal basis for  $H$  consisting of eigenvectors of  $T$ , i.e. a set  $\{e_j\}$  such that every  $x \in H$  can be written*

$$x = \sum_j (x, e_j)e_j.$$

*Proof* By Theorem 4.5 there exists a  $w_1$  such that  $\|w_1\| = 1$  and  $Tw_1 = \pm\|T\|w_1$ . Consider the subspace of  $H$  perpendicular to  $w_1$ ,

$$H_2 = w_1^\perp.$$

Then since  $T$  is self-adjoint, Lemma 4.13 shows that  $T$  leaves  $H_2$  invariant. If we consider  $T_2 = T|_{H_2}$  then we have  $T_2 \in B(H_2, H_2)$  with  $T_2$  compact; this operator is still self-adjoint, since for all  $x, y \in H_2$

$$(x, T_2y) = (x, Ty) = (T^*x, y) = (Tx, y) = (T_2x, y).$$

Now apply Theorem 4.5 to the operator  $T_2$  on the Hilbert space  $H_2$  find an eigenvalue  $\lambda_2 = \pm\|T_2\|$  and an eigenvector  $w_2 \in H_2$  with  $\|w_2\| = 1$ . Continue this process as long as  $T_n \neq 0$ .

If  $T_n = 0$  for some  $n$  then for any  $x \in H$  we have

$$y := x - \sum_{j=1}^{n-1} (x, w_j)w_j \in H_n.$$

Then

$$0 = T_n y = Ty = Tx - \sum_{j=1}^{n-1} (x, w_j)Tw_j = Tx - \sum_{j=1}^{n-1} \lambda_j(x, w_j)w_j$$

which is (4.2).

If  $T_n$  is never zero then consider

$$y_n := x - \sum_{j=1}^{n-1} (x, w_j) w_j \in H_n.$$

Then we have

$$\|x\|^2 = \|y_n\|^2 + \sum_{j=1}^{n-1} |(x, w_j)|^2,$$

and so  $\|y_n\| \leq \|x\|$ . It follows that

$$\left\| Tx - \sum_{j=1}^{n-1} \lambda_j (x, w_j) w_j \right\| = \|Ty_n\| \leq \|T_n\| \|y_n\| = |\lambda_n| \|x\|,$$

and since  $|\lambda_n| \rightarrow 0$  as  $n \rightarrow \infty$  we have (4.2).

Finally, let  $\{e_j\}$  be an orthonormal basis for  $\text{Ker } T$ ; each  $e_j$  is an eigenvector of  $T$  with eigenvalue zero, and since  $Te_j = 0$  but  $Tw_j = \lambda_j w_j$  with  $\lambda_j \neq 0$ , we know that  $(w_j, e_k) = 0$  for all  $j, k$ . So  $\{w_j\} \cup \{e_k\}$  is a countable orthonormal set in  $H$ .

Now, (??) implies that

$$T \left[ x - \sum_{j=1}^{\infty} (x, w_j) w_j \right] = 0,$$

i.e. that  $x - \sum_{j=1}^{\infty} (x, w_j) w_j \in \text{Ker } T$ , and therefore

$$x - \sum_{j=1}^{\infty} (x, w_j) w_j = \sum_{k=1}^{\infty} \alpha_k e_k,$$

since  $\{e_k\}$  is a basis for  $\text{Ker } T$ . It follows that  $\{w_j\} \cup \{e_k\}$  is an orthonormal basis for  $H$ .  $\square$

Note that it follows that when  $\text{Ker } T$  is trivial, i.e. when  $T$  is invertible, that the eigenvectors corresponding to non-zero eigenvalues of  $T$  span  $H$ .

#### 4.5 Eigenvalues and eigenfunctions for elliptic PDEs

Consider the operator

$$Lu = - \sum_{i,j=1}^n \partial_i (a_{ij}(x) \partial_j u), \quad (4.3)$$

where  $a_{ij} = a_{ji}$ , and the weak form of the equation  $Lu = f$ ,

$$B(u, v) = \sum_{i,j} \int_{\Omega} a_{ij}(x) (\partial_i u) (\partial_i v) dx = \int f v dx = (f, v) \quad \text{for all } v \in H_0^1(\Omega).$$

We know that  $B$  is symmetric.

We have already observe that the solution mapping  $u = Tf$  is compact. To show that it is symmetric, take  $f, g \in L^2(\Omega)$  and let  $u = Tf$  and  $v = Tg$ . Then  $u, v \in H_0^1(\Omega)$  are both allowable ‘test functions’, and so

$$(f, Tg) = \underbrace{(f, v)}_{Lu=f} = B(u, v) = \underbrace{B(v, u)}_{Lv=g} = (g, u) = (g, Tf).$$

**Corollary 4.15** *If  $L$  is a second order elliptic operator (i.e. (3.4) is satisfied) as in (4.3) with  $a_{ij} = a_{ji}$  then there exists a finite or countably infinite sequence of  $C^\infty(\Omega)$  eigenfunctions  $\{u_n\}$  satisfying*

$$Lu_n = \lambda_n u_n,$$

where  $\lambda_n \rightarrow \infty$  as  $n \rightarrow \infty$ . These eigenfunctions form an orthonormal basis for  $L^2(\Omega)$ .

*Proof* We know that  $T$  is a compact symmetric operator from  $L^2(\Omega)$  into itself; let us show in addition that  $\text{Ker } T = \{0\}$ . Suppose that  $f \in \text{Ker } T$ ; this means that  $Tf = 0$ , i.e. that

$$0 = B(0, v) = (f, v) \quad \text{for all } v \in H_0^1(\Omega),$$

which implies that  $f = 0$ <sup>1</sup>. It follows that  $T$  has a countable set of eigenfunctions  $u_n$  with  $Tu_n = \lambda_n u_n$ , where  $\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ , which form an orthonormal basis for  $L^2(\Omega)$ .

Since  $Tu_n$  is the solution of the weak form of the equation for  $f = u_n$ , we have

$$B(\lambda_n u_n, v) = (u_n, v) \quad \text{for all } v \in H_0^1(\Omega),$$

so that

$$B(u_n, v) = \left( \frac{u_n}{\lambda_n}, v \right) \quad \text{for all } v \in H_0^1(\Omega);$$

set  $\mu_n = \lambda_n^{-1}$ . Corollary 3.5 shows that  $u \in C^\infty(\Omega)$ . □

<sup>1</sup> This is not entirely straightforward, since  $f \in L^2(\Omega)$  but we have to take  $v \in H_0^1(\Omega)$ ; see examples.

## 5

---

### Linear parabolic problems

We now want to consider linear parabolic problems, the model of which is

$$\frac{\partial u}{\partial t} - \Delta u = f(x, t) \quad u|_{\partial\Omega} = 0 \quad (5.1)$$

subject to the initial condition

$$u(x, 0) = u_0(x).$$

What sort of solution do we expect? Assuming that  $u$  is smooth, if we multiply by  $u$  and integrate over  $\Omega$  then we obtain

$$\int_{\Omega} u_t(x, t)u(x, t) \, dx + \int_{\Omega} |\nabla u(x, t)|^2 \, dx = \int_{\Omega} f(x, t)u(x, t) \, dx,$$

so that

$$\frac{1}{2} \int_{\Omega} |u(x, t)|^2 \, dx + \int_{\Omega} |\nabla u(x, t)|^2 \, dx \leq \|f(\cdot, t)\|_{L^2(\Omega)} \|u(\cdot, t)\|_{L^2(\Omega)},$$

i.e.

$$\frac{1}{2} \frac{d}{dt} \|u(\cdot, t)\|_{L^2(\Omega)}^2 + \|\nabla u(\cdot, t)\|_{L^2(\Omega)}^2 \leq c \|f(\cdot, t)\|_{L^2(\Omega)} \|\nabla u(\cdot, t)\|_{L^2(\Omega)},$$

using Poincaré's inequality ( $\|u\|_{L^2} \leq c \|\nabla u\|_{L^2}$ ). Now we can use Young's inequality ( $2ab \leq a^2 + b^2$ ) on the right-hand side to obtain

$$\frac{d}{dt} \|u(\cdot, t)\|_{L^2(\Omega)}^2 + \|\nabla u(\cdot, t)\|_{L^2(\Omega)}^2 \leq c^2 \|f(\cdot, t)\|_{L^2(\Omega)}^2.$$

Integrating this equation from 0 to  $t$  gives

$$\|u(\cdot, t)\|_{L^2(\Omega)}^2 + \int_0^t \|\nabla u(\cdot, s)\|_{L^2(\Omega)}^2 \, ds \leq \|u_0(\cdot)\|_{L^2(\Omega)}^2 + c^2 \int_0^t \|f(\cdot, s)\|_{L^2(\Omega)}^2 \, ds.$$

This tells us that if  $u_0 \in L^2(\Omega)$  and

$$\int_0^T \|f(\cdot, s)\|_{L^2(\Omega)}^2 ds < \infty$$

(we will weaken this second condition), then we should expect a solution  $u(x, t)$  for which  $u(x, t) \in L^2(\Omega)$  for every  $t \geq 0$ ,

$$\sup_{0 \leq t \leq T} \|u(\cdot, t)\|_{L^2(\Omega)}^2 \leq M := \|u_0(\cdot)\|_{L^2(\Omega)}^2 + c^2 \int_0^T \|f(\cdot, s)\|_{L^2(\Omega)}^2 ds$$

and such that the  $H^1$  norm of  $u$  is square integrable in time,

$$\int_0^t \|\nabla u(\cdot, s)\|_{L^2(\Omega)}^2 ds \leq M.$$

Writing this more compactly, we expect  $u \in L^\infty(0, T; L^2) \cap L^2(0, T; H^1)$ . Moreover, since

$$u_t = \Delta u + f(x, t),$$

and

$$|(\Delta u, v)| = |(\nabla u, \nabla v)| \leq \|\nabla u\|_{L^2} \|\nabla v\|_{L^2},$$

implies that  $\|\Delta u\|_{H^{-1}} \leq \|\nabla u\|_{L^2}$ , it follows that  $u_t \in L^2(0, T; H^{-1})$ .

A weak solution of (5.1) will be a solution  $u \in L^2(0, T; H_0^1)$  with  $u_t \in L^2(0, T; H^{-1})$ ; as we will see, other properties of  $u$  (like  $u \in C^0([0, T]; L^2)$ ) will follow from this.

### 5.1 Banach-space valued function spaces and ‘Sobolev-type’ results

Notation becomes much easier if we suppress the  $x$  dependence, i.e. we think of a solution  $u(x, t)$  as a map  $u : [0, T] \rightarrow X$ , where  $X$  is some space of functions on  $\Omega$ . Then the fact that  $u(t) \in X$  expresses the fact that for each  $t$ ,  $u(t)$  depends on  $X$ .

We now make the  $L^p(0, T; X)$  notation more explicit. Let  $X$  be a Banach space. We say that a function  $f : (0, T) \rightarrow X$  is in  $L^p(0, T; X)$ ,  $1 \leq p < \infty$ , if the map  $t \mapsto \|f(t)\|_X$  is measurable and

$$\int_0^T \|f(t)\|_X^p dt < \infty;$$

the norm in  $L^p(0, T; X)$  is

$$\|f\|_{L^p(0, T; X)} = \left( \int_0^T \|f(t)\|_X dt \right)^{1/p}.$$

We have  $f \in L^\infty(0, T; X)$  is the map  $t \mapsto \|f(t)\|_X$  is measurable and the

$$\|f\|_{L^\infty(0, T; X)} = \text{ess sup}_{0 \leq t \leq T} \|f(t)\|_X$$

is finite.

If  $X$  is complete then  $L^p(0, T; X)$  is complete. If  $X$  is a Hilbert space then  $L^2(0, T; X)$  is a Hilbert space. (See examples.)

The space  $C^0([0, T]; X)$  consists of all functions that are continuous into  $X$ , i.e. for which

$$\|f(t) - f(s)\|_X \rightarrow 0 \quad \text{as} \quad s \rightarrow t.$$

Just as we had to deal with weak spatial derivatives, we now also have to deal with weak time derivatives: if  $u \in L^1_{\text{loc}}(0, T; X)$  then  $u_t \in L^1_{\text{loc}}(0, T; X)$  is its weak time derivative if

$$\int_0^T u_t \varphi dt = - \int_0^T u \varphi_t dt$$

for every  $\varphi \in C_c^\infty(0, T)$ .

This is a little bit naughty, since we haven't discussed integration of Banach-space valued functions; the definition of  $L^1(0, T; X)$  only involved integrating *norms* of  $f(t)$ . If  $X$  is reflexive (see later) then provided that  $f$  is integrable (can be approximated by step functions) and  $f \in L^1(0, T; X)$  one can define the integral of  $f$ . Assuming that

$$\left\| \int_0^T f(t) dt \right\|_X \leq \int_0^T \|f(t)\|_X dt,$$

we can readily prove the following useful lemma.

**Lemma 5.1** *Suppose that  $u_n$  has weak time derivative  $\dot{u}_n$ , and that*

$$u_n \rightarrow u \quad \text{and} \quad \dot{u}_n \rightarrow v \quad \text{in} \quad L^1(0, T; X).$$

*Then  $v = \dot{u}$ .*

*Proof* Taking  $\varphi \in C_c^\infty(0, T)$ , we have

$$\int_0^T \dot{u}_n \varphi \, dt = - \int_0^T u_n \varphi_t \, dt. \quad (5.2)$$

Consider the left-hand side. We have

$$\begin{aligned} \left\| \int_0^T (\dot{u}_n - v) \varphi \, dt \right\|_X &\leq \int_0^T \|(\dot{u}_n - v)\varphi(t)\|_X \, dt \\ &\leq \int_0^T |\varphi(t)| \|\dot{u}_n - v\|_X \, dt \\ &\leq \|\varphi\|_{L^\infty(0, T)} \int_0^T \|\dot{u}_n - v\|_X \, dt, \end{aligned}$$

and so the left-hand side of the equality (5.2) tends to

$$\int v \varphi \, dt.$$

Similarly, the right-hand side tends to

$$- \int u \varphi_t \, dt,$$

whence  $v = \dot{u}$  as claimed.  $\square$

We now want to prove some ‘Sobolev’ type results for Banach-space valued function spaces. The first is a generalisation of  $H^1(0, T) \subset C^0([0, T])$ .

**Lemma 5.2** *Suppose that  $f \in H^1(0, T; X)$ , i.e. that  $f \in L^2(0, T; X)$  and  $\dot{f} \in L^2(0, T; X)$ . Then*

$$f(t) = f(s) + \int_s^t \dot{f}(\tau) \, d\tau \quad \text{for all } 0 \leq s \leq t \leq T, \quad (5.3)$$

we have  $f \in C^0([0, T]; X)$ , and

$$\sup_{0 \leq t \leq T} \|f(t)\|_X \leq C \|f\|_{H^1(0, T; X)}. \quad (5.4)$$

If you look carefully at the proof,  $f, \dot{f} \in L^1(0, T; X)$  is sufficient for all these results.

*Proof* Extend  $u$  to be zero outside  $[0, T]$ , and then mollify with respect to  $t$  to give a function  $u_h : [0, T] \rightarrow X$  that is smooth in  $t$ , and is such that

$$u_h, \dot{u}_h \rightarrow u, \dot{u} \quad \text{in } L^2(0, T; X).$$

Now, for  $h > 0$  we have

$$u_h(t) = u_h(s) + \int_s^t \dot{u}_h(\tau) d\tau,$$

and a variant of the argument of Lemma 5.1 gives (5.3). That  $f(t)$  is continuous into  $X$  follows since the integral of an  $L^1$  function is continuous.

For the norm bound, observe that it follows from (5.3) that

$$\begin{aligned} \|f(t)\|_X &\leq \|f(0)\|_X + \int_0^t \|\dot{f}(\tau)\|_X d\tau \\ &\leq \|f(0)\|_X + \left( \int_0^t d\tau \right)^{1/2} \left( \int_0^t \|\dot{f}(\tau)\|_X^2 d\tau \right)^{1/2} \\ &= \|f(0)\|_X + T^{1/2} \|\dot{f}\|_{L^2(0,T;X)}, \end{aligned}$$

and also

$$\|f(0)\|_X \leq \|f(t)\|_X + \int_0^t \|\dot{f}(\tau)\|_X d\tau \leq \|f(t)\|_X + T^{1/2} \|\dot{f}\|_{L^2(0,T;X)}.$$

Integrating this inequality with respect to  $t$  from 0 to  $T$ , to obtain

$$\begin{aligned} \|f(0)\|_X &\leq \frac{1}{T} \int_0^T \|f(t)\|_X dt + T^{1/2} \|\dot{f}\|_{L^2(0,T;X)} \\ &\leq \frac{1}{T^{1/2}} \left( \int_0^T \|f(t)\|_X dt \right)^{1/2} + T^{1/2} \|\dot{f}\|_{L^2(0,T;X)}, \end{aligned}$$

whence (5.4) follows.  $\square$

In the light of our heuristic discussion of weak solutions, the following lemma will be more useful. We use  $\langle f, g \rangle$  to denote the pairing of  $f \in H^{-1}$  with  $g \in H_0^1$  (we have previously written this  $f(g)$ ).

**Lemma 5.3** *Suppose that  $f \in L^2(0, T; H_0^1)$  and  $\dot{f} \in L^2(0, T; H^{-1})$ . Then*

(i)  $f \in C^0([0, T]; L^2)$ ,

(ii)

$$\frac{d}{dt} \|f(t)\|_{L^2}^2 = 2 \langle \dot{f}, f \rangle,$$

and

(iii)

$$\sup_{0 \leq t \leq T} \|f(t)\| \leq C \left( \|f\|_{L^2(0,T;H_0^1)} + \|\dot{f}\|_{L^2(0,T;H^{-1})} \right).$$

Note that (iii) shows that if  $f_n \rightarrow f$  in  $L^2(0, T; H_0^1)$ ,  $\dot{f}_n \rightarrow \dot{f}$  in  $L^2(0, T; H^{-1})$ , then in fact  $f_n \rightarrow f$  in  $C^0([0, T]; L^2)$  (i.e.  $f_n(t) \rightarrow f(t)$  in  $L^2$ , uniformly for all  $t \in [0, T]$ .)

*Proof* Extend  $f$  by zero outside  $[0, T]$  and mollify to give  $f_h$  such that

$$f_h \rightarrow f \text{ in } L^2(0, T; H_0^1) \quad \text{and} \quad \dot{f}_h \rightarrow \dot{f} \text{ in } L^2(0, T; H^{-1}).$$

Then for any  $h > 0$  and any  $t^* \in [0, T]$ , since  $f_h$  is smooth in time

$$\|f_h(t)\|_{L^2}^2 = \|f_h(t^*)\|_{L^2}^2 + 2 \int_{t^*}^t \langle \dot{f}_h(s), f_h(s) \rangle \, ds.$$

Now choose  $t^* \in [0, T]$  such that

$$\|f_h(t^*)\|_{L^2}^2 = \frac{1}{T} \int_0^T \|f_h(t)\|_{L^2}^2 \, dt.$$

Then

$$\|f_h(t)\|_{L^2}^2 \leq \frac{1}{T} \int_0^T \|f_h(t)\|_{L^2}^2 \, dt + \int_0^T \|\dot{f}_h(s)\|_{H^{-1}} \|f_h(s)\|_{H^1} \, ds,$$

i.e.

$$\sup_{0 \leq t \leq T} \|u_h(t)\|_{L^2}^2 \leq C \left( \|f_h\|_{L^2(0, T; H^1)}^2 + \|\dot{f}_h\|_{L^2(0, T; H^{-1})}^2 \right).$$

Letting  $h \rightarrow 0$  gives (iii), (ii), and (i).  $\square$

## 5.2 Weak solutions of parabolic problems

Assume that  $u$  is smooth, and take the inner product of (5.1) with some fixed  $v \in C_c^\infty(\Omega)$ ; then for each fixed  $t$  we obtain

$$(u_t, v) - (\Delta u, v) = (f(t), v),$$

and integrating by parts we obtain

$$(u_t, v) + (\nabla u, \nabla v) = (f(t), v). \quad (5.5)$$

Arguing as before, using the density of  $C_c^\infty(\Omega)$  in  $H_0^1(\Omega)$ , we expect (5.5) to hold for every  $v \in H_0^1(\Omega)$ . We have made the equation ‘weak in space’. We make it ‘weak in time’ by allowing  $u_t$  to be the weak time derivative  $\dot{u}$ .

We have therefore reformulated the problem as: given  $u_0 \in L^2(\Omega)$  and

$f(t) \in L^2(0, T; H^{-1}(\Omega))$ , find  $u \in L^2(0, T; H_0^1(\Omega))$  with  $\dot{u} \in L^2(0, T; H^{-1}(\Omega))$  such that  $u(0) = u_0$ , and

$$(\dot{u}, v) + (\nabla u, \nabla v) = \langle f(t), v \rangle \quad \text{for all } v \in H_0^1(\Omega)$$

and for almost every  $t \in (0, T)$  (since changing a function on a set of measure zero will not affect its weak derivative).

This formulation is equivalent to the following, which will prove useful in our existence/uniqueness proof: find  $u$  with the regularity above such that

$$\dot{u} - \Delta u = f(t)$$

as an equality in  $L^2(0, T; H^{-1})$ .

### 5.3 Galerkin approximations

We show the existence and uniqueness of weak solutions using the Galerkin method – we approximate the original infinite-dimensional problem by a sequence of finite-dimensional problems, and then take the limit. In order to do this we will use the orthonormal basis of  $L^2(\Omega)$  given by the eigenfunctions  $\{w_j\}_{j=1}^\infty$  of the Laplacian on  $\Omega$  with Dirichlet boundary conditions (as in Corollary 4.15). In fact these eigenfunctions are also an orthogonal basis for  $H_0^1(\Omega)$ ; they are a basis since  $L^2(\Omega) \supset H_0^1(\Omega)$ , and they are orthogonal since

$$\begin{aligned} (w_k, w_j)_{H^1} &= (w_k, w_j)_{L^2} + (\nabla w_k, \nabla w_j)_{L^2} = \delta_{jk} + (w_k, (-\Delta)w_j) \\ &= \delta_j + \lambda_j(w_k, w_k) = (1 + \lambda_j)\delta_{jk}. \end{aligned}$$

Now, if we take  $u \in L^2(\Omega)$  then we can approximate  $u$  in the space spanned by the first  $n$  eigenfunctions,

$$P_n u = \sum_{j=1}^n (u, w_j) w_j.$$

The same definition works for  $u \in H_0^1(\Omega)$  (we still take the inner product in  $L^2$ ), and for  $f \in H^{-1}(\Omega)$  we define  $P_n f$  by

$$\langle P_n f, u \rangle = \langle f, P_n u \rangle \quad \text{for all } u \in H_0^1(\Omega).$$

Note that for  $u, v \in L^2$ ,

$$(u, P_n v) = (P_n u, v) = (P_n u, P_n v)$$

and similarly in  $H_0^1$  (see examples).

We now have the following.

**Lemma 5.4** *Let  $X = H_0^1(\Omega)$ ,  $L^2(\Omega)$ , or  $H^{-1}(\Omega)$ , and take  $f \in X$ . Then*

$$\|P_n f\|_X \leq \|f\|_X \quad \text{and} \quad P_n f \rightarrow f \text{ in } X.$$

*Proof* If  $X = L^2$  then since  $\{w_j\}_{j=1}^\infty$  is an orthonormal basis,

$$f = \sum_{j=1}^{\infty} (f, w_j) w_j \quad \text{and} \quad \|f\|_{L^2}^2 = \sum_{j=1}^{\infty} |(f, w_j)|^2,$$

which implies the result. For  $X = H_0^1(\Omega)$ , we have

$$(\nabla f, \nabla f) = (f, -\Delta f) = \sum_{j=1}^{\infty} \lambda_j |(f, w_j)|^2,$$

and again the result follows. For  $X = H^{-1}(\Omega)$ , for the norm bound we have

$$|\langle P_n f, v \rangle| = |\langle f, P_n v \rangle| \leq \|f\|_{H^{-1}} \|P_n v\|_{H_0^1} \leq \|f\|_{H^{-1}} \|v\|_{H_0^1},$$

so that  $\|P_n f\|_{H^{-1}} \leq \|f\|_{H^{-1}}$ . For the convergence, use the Riesz Representation Theorem to find a  $\varphi \in H_0^1$  such that

$$\langle f, v \rangle = (\varphi, v)_{H_0^1} \quad \text{for all } v \in H_0^1.$$

Then

$$\|P_n f - f\|_{H^{-1}} = \sup_{v \in H_0^1: \|v\|_{H_0^1}=1} |\langle P_n f - f, v \rangle|,$$

and for  $\|v\|_{H_0^1} = 1$ ,

$$\begin{aligned} |\langle P_n f - f, v \rangle| &= |\langle f, P_n v - v \rangle| \\ &= |(\varphi, P_n v - v)_{H_0^1}| \\ &= |(P_n \varphi - \varphi, v)_{H_0^1}| \\ &\leq \|P_n \varphi - \varphi\|_{H^1} \|v\|_{H^1} \\ &= \|P_n \varphi - \varphi\|_{H^1}, \end{aligned}$$

i.e.  $\|P_n f - f\|_{H^{-1}} \rightarrow 0$  as required.  $\square$

We now find approximate solutions  $u_n(t)$  contained in the linear span of  $\{w_1, \dots, w_n\}$ ,

$$u_n(t) = \sum_{j=1}^n u_{nj}(t)w_j.$$

Note that since each  $w_j$  is a smooth function of  $x$ , each of the  $u_n$  is a smooth function of  $x$  (for each fixed  $t$ ). So we can integrate by parts, etc., rigorously.

We have two ways of viewing the equations for  $u_n$ , which are

$$(\dot{u}_n, w_j) + B(u_n, w_j) = \langle f(t), w_j \rangle \quad j = 1, \dots, n$$

First, as a set of  $n$  ODEs for the coefficients  $u_{nj}$ , which form a vector  $\underline{u}_n \in \mathbb{R}^n$ ,

$$\dot{u}_{nj} + \lambda_j u_{nj} = f_j(t) := \langle f(t), w_j \rangle, \quad (5.6)$$

or as a PDE for the (smooth) function  $u_n$ ,

$$\frac{\partial u_n}{\partial t} - \Delta u_n = P_n f(t). \quad (5.7)$$

The initial condition is  $u_n(0) = P_n u_0$ .

Using standard theory for ODES, the  $n$  ODEs in (5.6) have - at least for a short time - a unique solution  $u_n(t)$ . Problems with this solution only arise if the solutions ‘blows up’, i.e. if its norm tends to infinity. Since

$$\|\underline{u}_n\|_{\mathbb{R}^n}^2 = \sum_{j=1}^n |u_{nj}|^2 = \|u_n\|_{L^2(\Omega)}^2,$$

we could also show that the  $L^2$  norm of the function  $u_n$  remains bounded. This is probably easier; in fact, we have already done these calculations in our ‘heuristic’ analysis: take the inner product of (5.7) with  $u_n$  and integrate by parts (we do this because  $u_n$  is smooth), so that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u_n\|_{L^2}^2 + \|\nabla u_n\|^2 &= \langle P_n f(t), u_n \rangle \\ &= \langle f(t), P_n u_n \rangle = \langle f(t), u_n \rangle \\ &\leq \|f(t)\|_{H^{-1}} \|u_n\|_{H^1} \\ &\leq \|f(t)\|_{H^{-1}} \|\nabla u_n\|_{L^2}. \end{aligned}$$

Using  $2ab \leq a^2 + b^2$  we obtain

$$\frac{d}{dt} \|u_n\|_{L^2}^2 + \|\nabla u_n\|^2 \leq c^2 \|f(t)\|_{H^{-1}}^2,$$

and hence, integrating from 0 to  $t$ ,

$$\begin{aligned} \|u_n(t)\|_{L^2}^2 + \int_0^t \|\nabla u_n(s)\|_{L^2}^2 ds &\leq \|u_n(0)\|_{L^2}^2 + c^2 \int_0^t \|f(s)\|_{H^{-1}}^2 ds \\ &\leq \|u_0\|_{L^2}^2 + c^2 \int_0^T \|f(s)\|_{H^{-1}}^2 ds, \end{aligned}$$

for all  $t \in [0, T]$ .

So we have an approximate solution  $u_n$  that is bounded in  $L^\infty(0, T; L^2)$  and  $L^2(0, T; H_0^1)$  with the bound independent of  $n$ . Arguing as we did before, we also have a bound on  $\dot{u}_n$  in  $L^2(0, T; H^{-1})$  that does not depend on  $n$ .

We now need to show that these approximate solutions in fact converge to some  $u$  as  $n \rightarrow \infty$ , and that this  $u$  satisfies the equation. These are separate issues; it is possible to have  $u_n$  converges to  $u$  but not in a strong enough sense that  $u$  actually satisfies the limiting equation (this will, fortunately, not be the case here).

In this simple linear case we can show the convergence directly: consider

$$\frac{du_n}{dt} - \Delta u_n = P_n f(t) \quad u_n(0) = P_n u_0$$

and

$$\frac{du_m}{dt} - \Delta u_m = P_m f(t) \quad u_m(0) = P_m u_0.$$

We want to show that  $\{u_n\}$  (and also  $\{\dot{u}_n\}$ ) is Cauchy. To do this, consider  $w_{nm} = u_n - u_m$ , with  $n > m$ . Then  $w_{nm}$  satisfies

$$\frac{dw_{nm}}{dt} - \Delta w_{nm} = (P_n - P_m)f(t) \quad w_{nm}(0) = (P_n - P_m)u_0.$$

If we repeat the calculations above we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|w_{nm}\|_{L^2}^2 + \|\nabla w_{nm}\|_{L^2}^2 &= \langle (P_n - P_m)f(t), w_{nm} \rangle \\ &\leq c \|(P_n - P_m)f(t)\|_{H^{-1}} \|\nabla w_{nm}\|_{L^2}, \end{aligned}$$

and so

$$\frac{d}{dt} \|w_{nm}\|_{L^2}^2 + \|\nabla w_{nm}\|_{L^2}^2 \leq \|(P_n - P_m)f(t)\|_{H^{-1}}^2.$$

Integrating from 0 to  $T$  we obtain

$$\|w_{nm}(T)\|_{L^2}^2 + \int_0^T \|\nabla w_{nm}(s)\|_{L^2}^2 ds \leq \|w_{nm}(0)\|_{L^2}^2 + \int_0^T \|(P_n - P_m)f(t)\|_{H^{-1}}^2 dt.$$

Choose  $\epsilon > 0$ . Then we can make  $\|w_{nm}(0)\|_{L^2}^2$  as small as we wish, since  $P_n u_0 \rightarrow u_0$  in  $L^2$ . We show that  $P_n f \rightarrow f$  in  $L^2(0, T; H^{-1})$ , from which it follows that  $\{P_n f\}$  is Cauchy in the same space, and hence that  $w_{nm}$  is Cauchy in  $L^2(0, T; H^1)$ . Consider

$$\int_0^T \|P_n f(t) - f(t)\|_{H^{-1}}^2 dt. \quad (5.8)$$

We know that  $P_n f(t) \rightarrow f(t)$  in  $H^{-1}$  for each fixed  $t$ , i.e. that  $\|P_n f(t) - f(t)\|_{H^{-1}}^2 \rightarrow 0$  for each fixed  $t$ ; and we also know that

$$\|P_n f(t) - f(t)\|_{H^{-1}}^2 \leq (\|P_n f(t)\|_{H^{-1}} + \|f(t)\|_{H^1})^2 \leq 4\|f(t)\|_{H^{-1}}^2 \in L^1(0, T).$$

So we can use the Dominated Convergence Theorem to show that (5.8) converges to zero.

So  $u_n$  converges to some  $u$  in  $L^2(0, T; H^1)$  (and also in  $L^\infty(0, T; L^2)$ , but we will soon do better than this). It follows that  $-\Delta u_n \rightarrow -\Delta u$  in  $L^2(0, T; H^{-1})$ ; since we have already seen that  $P_n f \rightarrow f$  in  $L^2(0, T; H^{-1})$ , it follows from the equation

$$\dot{u}_n = -\Delta u_n + P_n f(t)$$

that  $\dot{u}_n \rightarrow v$  in  $L^2(0, T; H^{-1})$ . Since  $u_n \rightarrow u$  in  $L^2(0, T; H^1)$  implies that  $u_n \rightarrow u$  in  $L^2(0, T; H^{-1})$ , we can use Lemma 5.1 to deduce that  $v = \dot{u}$ .

We have therefore shown that every term in the equation converges to its appropriate limit in  $L^2(0, T; H^{-1})$ , so  $u \in L^2(0, T; H_0^1)$ ,  $\dot{u} \in L^2(0, T; H^{-1})$ , and  $u$  satisfies

$$\dot{u} - \Delta u = f(t)$$

as an equality in  $L^2(0, T; H^{-1})$ . We remarked already that this was equivalent to the weak form (5.6).

All that is left is to check that  $u$  satisfies the initial condition and is unique. Since  $u_n$  is Cauchy in  $L^2(0, T; H^1)$  and  $\dot{u}_n$  is Cauchy in  $L^2(0, T; H^{-1})$ , it follows from Lemma 5.3 that in fact  $u_n$  is Cauchy in  $C^0([0, T]; L^2)$ . In particular,  $u(0) = \lim_{n \rightarrow \infty} u_n(0) = \lim_{n \rightarrow \infty} P_n u_0 = u_0$ , as required.

Finally, for uniqueness we again use Lemma 5.3, which ensure that if

$$\frac{du}{dt} - \Delta u = f(t) \quad \text{and} \quad \frac{dv}{dt} - \Delta v = f(t)$$

with  $u(0) = v(0)$  then  $\dot{w} = \Delta w$  and

$$\frac{1}{2} \frac{d}{dt} \|w\|_{L^2}^2 = \langle \dot{w}, w \rangle = \langle \Delta w, w \rangle = -\|\nabla w\|_{L^2}^2 \leq 0,$$

whence  $\|w(t)\|_{L^2}^2 = 0$  for all  $t \geq 0$ , i.e. the solutions are unique.

Combining this, we have proved the following:

**Theorem 5.5** *Given  $u_0 \in L^2$  and  $f \in L^2(0, T; H^{-1})$ , the equation*

$$u_t - \Delta u = f(t) \quad u|_{\partial\Omega} = 0 \quad u(0) = u_0$$

*has a unique weak solution  $u$ , with  $u \in L^2(0, T; H_0^1) \cap C^0([0, T]; L^2)$  and  $\dot{u} \in L^2(0, T; H^{-1})$ , such that  $u(0) = u_0$  and*

$$\dot{u} - \Delta u = f(t)$$

*as an equality in  $L^2(0, T; H^{-1})$ .*

One can also obtain higher regularity of solutions assuming higher regularity of  $u_0$  and  $f$ , for example: if  $u_0 \in H_0^1(\Omega)$  and  $f \in L^2(0, T; L^2)$  then in fact  $u \in L^2(0, T; H^2) \cap C^0([0, T]; H_0^1)$  and  $\dot{u} \in L^2(0, T; L^2)$ . (The equation then holds as an equality in  $L^2(0, T; L^2)$ , and in particular almost everywhere in  $(0, T) \times \Omega$ ). To prove this rigorously one performs the following heuristic computations using the Galerkin approximations and then takes limits; but essentially the result follows from the following ‘formal estimates’ (‘formal’ here means ‘assume that everything is smooth enough to carry out any manipulations you like’). Take the inner product with  $-\Delta u$  to give

$$\int_{\Omega} u_t \cdot (-\Delta u) \, dx + \|\Delta u\|_{L^2}^2 = (f(t), -\Delta u) \leq \|f(t)\|_{L^2} \|\Delta u\|_{L^2}.$$

The first term is

$$\begin{aligned} - \sum_{i,j=1}^n \int_{\Omega} (\partial_t u_j) (\partial_i^2 u_j) &= \sum_{i,j=1}^n \int_{\Omega} \partial_t (\partial_i u_j) (\partial_i u_j) = \frac{1}{2} \frac{d}{dt} \sum_{i,j=1}^n \int_{\Omega} |\partial_i u_j|^2 \\ &= \frac{1}{2} \frac{d}{dt} \|\nabla u\|^2, \end{aligned}$$

so

$$\frac{1}{2} \frac{d}{dt} \|\nabla u\|^2 + \|\Delta u\|^2 \leq \|f(t)\|_{L^2} \|\Delta u\|_{L^2},$$

which using Young’s inequality gives

$$\frac{d}{dt} \|\nabla u\|^2 + \|\Delta u\|^2 \leq \|f(t)\|_{L^2}^2.$$

Integrating from 0 to  $t$  we obtain

$$\|\nabla u(t)\|_{L^2}^2 + \int_0^t \|\Delta u(s)\|_{L^2}^2 ds \leq \|\nabla u(0)\|_{L^2}^2 + \int_0^t \|f(s)\|_{L^2}^2 ds.$$

Since  $u_0 \in H_0^1$  and  $f \in L^2(0, T; L^2)$  this shows that we should expect  $u \in L^\infty(0, T; H_0^1)$  and

$$\int_0^t \|\Delta u(s)\|_{L^2}^2 ds < \infty.$$

Since the theory of elliptic regularity shows that when  $-\Delta u = f$  then  $\|u\|_{H^2} \leq c\|f\|_{L^2}$ , i.e. that

$$\|u\|_{H^2} \leq c\|\Delta u\|_{L^2}$$

it follows also that  $u \in L^2(0, T; H^2)$ . Now, if  $u \in H^2$ , then  $\Delta u \in L^2$ . So from  $\dot{u} = \Delta u + f(t)$  we expect  $\dot{u} \in L^2(0, T; L^2)$ .

We know that if  $f \in L^2(0, T; H_0^1)$  and  $\dot{f} \in L^2(0, T; H^1)$  then  $f \in C^0([0, T]; L^2)$ . One can in fact show - with just a little more work - that the same result hold if  $f \in L^2(0, T; H^1)$  ( $H^1$  now, not  $H_0^1$ ). Using this (which we haven't proved), one can show that the increased regularity of  $u$  and  $\dot{u}$  implies that  $u$  is continuous into  $H^1$ : indeed, if  $u \in L^2(0, T; H^2)$  then for any  $j = 1, \dots, n$ , we have  $\partial_j u \in L^2(0, T; H^1)$  and  $\partial_j \dot{u} \in L^2(0, T; H^{-1})$ . So then  $\partial_j u \in C^0([0, T]; L^2)$ , which shows that  $u \in C^0([0, T]; H^1)$ .

There are of course even higher regularity results, e.g. if  $u_0 \in H^k$  and  $f \in L^2(0, T; H^{k-1})$  then  $u \in L^2(0, T; H^{k+1})$ ,  $\dot{u} \in L^2(0, T; H^{k-1})$ , and  $u \in C^0([0, T]; H^k)$ . See Evans (\*\*\*\*) for details. One can also prove regularity in time with additional assumptions on  $f$ ; see the examples.

## 6

---

### Nonlinear equations, weak convergence

In our previous analysis we first obtained uniform estimates on the Galerkin approximations: we showed that  $u_n$  was bounded in  $L^\infty(0, T; L^2)$  and  $L^2(0, T; H_0^1)$ , and the  $\dot{u}_n$  was bounded in  $L^2(0, T; H^{-1})$ , with the bounds independent of  $n$ .

Because the equations were linear, we could then in fact show directly that  $\{u_n\}$  was Cauchy in these spaces. But more generally we cannot proceed directly: we require some compactness results to guarantee that boundedness of  $u_n$  allows us to find subsequences that converge in some sense. We now introduce such a notion of convergence, and prove the required compactness theorem.

#### 6.1 Dual spaces, the Hahn–Banach Theorem, and weak convergence

We have briefly introduced dual spaces earlier, and we now consider them in a little more detail.

Recall that if  $X$  is a Banach space, its dual space  $X^*$  consists of all bounded linear functionals from  $X$  into  $\mathbb{K}$  (although we restrict to the case  $\mathbb{K} = \mathbb{R}$  here), i.e.  $X^* = \mathcal{L}(X, \mathbb{K})$ . The norm on  $X^*$  is the  $\mathcal{L}(X, \mathbb{K})$  norm (‘operator norm’),

$$\|f\|_{X^*} = \sup_{x \in X: \|x\|_X=1} |f(x)|.$$

The following theorem is somehow the first and central result in the study

of dual spaces. It shows that a linear functional defined on a subspace of  $X$  can be extended to a linear functional defined on the whole of  $X$  without increasing its norm.

**Theorem 6.1 (Hahn–Banach Theorem)** *Let  $X$  be a Banach space, and  $U$  a subspace of  $X$ . Suppose that  $f : U \rightarrow \mathbb{R}$  is a linear functional on  $U$  such that*

$$|f(x)| \leq M\|x\| \quad \text{for all } x \in U.$$

*Then there exists an  $F \in X^*$  that extends  $f$  (i.e.  $F(x) = f(x)$  for all  $x \in U$ ) and does not increase its norm,*

$$|F(x)| \leq M\|x\| \quad \text{for all } x \in X.$$

For a simple proof in a Hilbert space see the examples. The result in full generality relies on Zorn's Lemma, i.e. on the Axiom of Choice. Although it looks fairly innocent, it has some interesting consequences (see examples).

As an immediate application, we prove the existence of a particularly useful class of linear functionals, and show therefore that understanding linear functionals is in some way enough to understand elements of  $X$ .

**Lemma 6.2** *Let  $x \in X$ . Then there exists an  $f \in X^*$  such that  $\|f\|_{X^*} = 1$  and  $f(x) = \|x\|$ .*

*Proof* Define  $\hat{f}$  on the linear space  $U$  spanned by  $x$  as

$$\hat{f}(\alpha x) = \alpha\|x\|.$$

Then  $\hat{f}(x) = \|x\|$  and  $|\hat{f}(z)| \leq \|z\|$  for all  $z \in U$ . Extend  $\hat{f}$  to an  $f \in X^*$ ; then  $\|f\|_{X^*} = 1$  and  $f(x) = \hat{f}(x) = \|x\|$ .  $\square$

**Corollary 6.3** *Let  $x, y \in X$ . If  $f(x) = f(y)$  for every  $f \in X^*$  then  $x = y$ .*

*Proof* If  $x \neq y$  then by the previous lemma there exists an  $f$  with  $\|f\|_{X^*} = 1$  such that  $f(x) - f(y) = f(x - y) = \|x - y\| \neq 0$ .  $\square$

We say that  $\{x_n\} \in X$  converges weakly to  $x \in X$ , and write  $x_n \rightharpoonup x$ , if

$$f(x_n) \rightarrow f(x) \quad \text{for all } f \in X^*.$$

If  $x_n \rightarrow x$  then  $x_n \rightharpoonup x$ ; for any  $f \in X^*$

$$|f(x_n) - f(x)| \leq \|f\|_{X^*} \|x_n - x\|_X \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Note that in a Hilbert space, where every linear functional is of the form  $(y, \cdot)$  for some  $y \in H$ ,  $x_n \rightharpoonup x$  if

$$(x_n, y) \rightarrow (x, y) \quad \text{for all } y \in H.$$

This provides an easy example of a sequence that converges weakly but does not converge; pick a countable orthonormal set  $\{e_j\}_{j=1}^\infty$ . Then for any  $y \in H$  Bessel's inequality

$$\sum_{j=1}^{\infty} |(y, e_j)|^2 \leq \|y\|^2$$

shows that the sum converges; it follows that  $(y, e_j) \rightarrow 0$  as  $j \rightarrow \infty$ , and hence that  $e_j \rightharpoonup 0$ . But the sequence  $\{e_j\}$  does not converge (any two elements are a distance  $\sqrt{2}$  apart).

**Lemma 6.4** *If  $x_n \rightharpoonup x$  then*

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|.$$

*Proof* Choose  $f \in X^*$  with  $\|f\|_{X^*} = 1$  such that  $f(x) = \|x\|$ . Then

$$\|x\| = f(x) = \lim_{n \rightarrow \infty} f(x_n),$$

so

$$\|x\| \leq \liminf_{n \rightarrow \infty} |f(x_n)| \leq \liminf_{n \rightarrow \infty} \|f\|_{X^*} \|x_n\|_X;$$

the result follows since  $\|f\|_{X^*} = 1$ .  $\square$

In fact any weakly convergent sequence is bounded; this is not straightforward (see examples).

**Lemma 6.5** *Weak limits are unique.*

*Proof* Suppose that  $x_n \rightharpoonup x$  and  $x_n \rightharpoonup y$ . Then for any  $f \in X^*$ ,  $f(x) = \lim_{n \rightarrow \infty} f(x_n) = f(y)$ . So by Lemma 6.3,  $x = y$ .  $\square$

The following result is often useful.

**Lemma 6.6** *Suppose that  $X, Y$  are Banach spaces, with  $Y$  compactly embedded in  $X$ . Then if  $y_n \rightharpoonup y$  in  $Y$ ,  $y_n \rightarrow y$  in  $X$ .*

*Proof* First, observe that  $y_n \rightharpoonup y$  in  $Y$ . Take  $f \in Y^*$ , then  $Y$  is a linear subspace of  $X$  so there is an  $F \in X^*$  such that  $F = f$  on  $Y$ , and so

$$f(y_n) = F(y_n) \rightarrow F(y) = f(y).$$

Now, if  $y_n \not\rightarrow y$  in  $X$ , then there is an  $\epsilon > 0$  and a subsequence (which we relabel)  $y_n$  such that  $\|y_n - y\|_X > \epsilon$  for every  $j$ . Since  $\{y_n\}$  is a bounded sequence in  $Y$ , and  $Y$  is compactly embedded in  $X$ , it has a subsequence (which we relabel) that converges to some  $z \in X$ . If  $y_n \rightarrow z$  in  $X$  then  $y_n \rightharpoonup z$  in  $X$ ; but weak limits are unique, so  $z = y$ , a contradiction.  $\square$

There is another notion of weak convergence, weak-\* convergence, which deals with sequences of elements of  $X^*$ . If  $\{f_n\} \in X^*$  then  $f_n$  converges weakly-\* to  $f$ ,  $f_n \xrightarrow{*} f$  if

$$f_n(x) \rightarrow f(x) \quad \text{for all } x \in X.$$

**Theorem 6.7** *Suppose that  $X$  is separable. Then a bounded sequence in  $X^*$  has a weakly-\* convergent subsequence.*

*Proof* Let  $\{x_k\}$  be a countable dense subset of  $X$ , and  $\{f_j\}$  a sequence in  $X^*$  such that  $\|f_j\|_{X^*} \leq M$ . A standard diagonal argument yields a subsequence of the  $\{f_j\}$  (which we relabel) such that  $f_j(x_k)$  converges for every  $k$ . To show that in fact  $f_j(x)$  converges for every  $x \in X$ , given  $\epsilon > 0$  find  $x_k$  such that  $\|x - x_k\| < \epsilon/3M$ , and write

$$\begin{aligned} |f_i(x) - f_j(x)| &\leq |f_i(x) - f_i(x_k)| + |f_i(x_k) - f_j(x_k)| + |f_j(x_k) - f_j(x)| \\ &\leq \|f_i\|_{X^*} \|x - x_k\| + |f_i(x_k) - f_j(x_k)| + \|f_j\|_{X^*} \|x_k - x\| \\ &\leq \frac{2\epsilon}{3} + |f_i(x_k) - f_j(x_k)|. \end{aligned}$$

Since  $\{f_j(x_k)\}$  is Cauchy, one can find an  $N$  such that the second term is  $< \epsilon/3$  for all  $i, j \geq N$ . One then shows as in the proof of Theorem 1.3 that  $f$  defined for each  $x \in X$  by  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$  is an element of  $X^*$  with  $\|f\|_{X^*} \leq M$ .  $\square$

We can convert this into a compactness result for bounded subsets of  $X$  under a condition that relates the second dual of  $X$ ,  $(X^*)^*$ , to  $X$  itself...

Since  $X^*$  is a Banach space (Theorem 1.3), we can consider its dual,  $(X^*)^*$  (usually written just  $X^{**}$ ). We already have a large collection of elements of  $X^{**}$  to hand; given any  $x \in X$ , define

$$L_x(f) = f(x) \quad \text{for all } f \in X^*.$$

Then  $L_x$  is linear in  $f$ , and since

$$|L_x(f)| = |f(x)| \leq \|f\|_{X^*} \|x\|_X$$

and if  $f \in X^*$  has  $f(x) = \|x\|$  and  $\|f\|_{X^*} = 1$  then

$$|L_x(f)| = |f(x)| = \|x\| = \|f\|_{X^*} \|x\|_X$$

it follows that  $\|L_x\|_{X^{**}} = \|x\|_X$ . The map  $x \mapsto L_x$  maps  $X$  isometrically onto a subspace of  $X^{**}$ . If this map is onto, i.e. if for every  $G \in X^{**}$  there exists an  $x$  such that  $G = L_x$ , then  $X$  is called *reflexive*. Loosely,  $X$  is reflexive if  $X = X^{**}$ ; but the equality here has to be understood in the sense that the two spaces are isometrically isomorphic under the map  $x \mapsto L_x$ .

Examples: any Hilbert space is reflexive (since  $H^* \simeq H$ ). Every  $L^p$  space with  $1 < p < \infty$  is reflexive, since  $(L^p)^* \simeq L^q$  with  $p, q$  conjugate: that every  $f \in L^q(\Omega)$  gives rise to a linear functional on  $L^p(\Omega)$  follows from Hölder's inequality: define

$$L_f(g) := \int_{\Omega} f(x)g(x) \, dx,$$

and then

$$|L_f(g)| \leq \|f\|_{L^q} \|g\|_{L^p}.$$

To show that  $\|L_f\|_{(L^p)^*} = \|f\|_{L^q}$ , observe that if  $f \in L^q$  then the function  $g = f|f|^{q-2} \in L^p$ , since for  $(p, q)$  conjugate  $p = q/(q-1)$ , so that

$$\|g\|_{L^p} = \left( \int_{\Omega} (|f|^{q-1})^p \right)^{1/p} = \left( \int_{\Omega} |f|^q \right) = \|f\|_{L^q}^{q/p}.$$

With this choice for  $g$ , we have

$$\begin{aligned} |L_f(g)| &= \left| \int_{\Omega} |f|^q \, dx \right| = \|f\|_{L^q}^q = \|f\|_{L^q} \|f\|_{L^q}^{q-1} = \|f\|_{L^q} \|g\|_{L^p}^{p(q-1)/q} \\ &= \|f\|_{L^q} \|g\|_{L^p}, \end{aligned}$$

which shows that  $\|L_f\|_{(L^p)^*} = \|f\|_{L^q}$  as required. Some more work is required to show that every such linear functional can be written as  $L_f$  for some  $f$ .

The dual of  $L^1$  is  $L^\infty$ ; the dual of  $L^\infty$  is *not*  $L^1$ . So  $L^1$  is not reflexive. Neither is  $L^\infty$ , since  $X$  is reflexive iff  $X^*$  is reflexive (a simple but head-spinning proof). (However,  $L^1$  is separable, so we can apply Theorem 6.7 to deduce that a sequence bounded in  $L^\infty = (L^1)^*$  has a weakly-\* convergent subsequence.)

We can now prove the following more comfortable compactness result.

**Theorem 6.8** *Let  $X$  be a reflexive Banach space with  $X^*$  separable. Then any bounded sequence in  $X$  has a weakly convergent subsequence.*

The result is still true (with a more complicated proof) dropping the requirement that  $X$  is separable.

*Proof* Take a bounded sequence  $\{x_n\} \in X$ . Consider the sequence  $G_n \in X^{**}$  with

$$G_n(f) = f(x_n) \quad \text{for all } f \in X^*.$$

Then  $\|G_n\|_{X^{**}} = \|x_n\|_X$ , so  $G_n$  is a bounded sequence in  $(X^*)^*$ . Theorem 6.7 implies that there is a subsequence that converges weakly-\* in  $(X^*)^*$  to some  $G \in X^{**}$ .

Since  $X$  is reflexive, there exists an  $x \in X$  such that  $G = L_x$ , i.e. such that

$$G(f) = f(x) \quad \text{for all } f \in X^*.$$

That  $G_n \xrightarrow{*} G$  in  $(X^*)^*$  says that

$$f(x_n) = G_n(f) \rightarrow G(f) = f(x) \quad \text{for all } f \in X^*.$$

This is precisely  $x_n \rightharpoonup x$ . □

## 6.2 Weak compactness and our linear parabolic PDE

If we return to our uniform bounds on the Galerkin solutions of our PDE

$$u_t - \Delta u = f(t)$$

we have  $u_n$  bounded in  $L^2(0, T; H_0^1)$  (also  $L^\infty(0, T; L^2)$ ) and  $\dot{u}_n$  bounded in  $L^2(0, T; H^{-1})$ .

Using the above compactness theorems, we can find subsequence such that

$$\begin{aligned} u_{n_j} &\rightharpoonup u && \text{in } L^2(0, T; H_0^1) \\ u_{n_j} &\overset{*}{\rightharpoonup} u && \text{in } L^\infty(0, T; L^2) \\ \dot{u}_{n_j} &\overset{*}{\rightharpoonup} v && \text{in } L^2(0, T; H^{-1}). \end{aligned}$$

Some comments are in order here:  $L^2(0, T; H_0^1)$  is a Hilbert space, hence reflexive, so we can use Theorem 6.8 to find a subsequence for which  $u_{n_j} \rightharpoonup u$  in  $L^2(0, T; H_0^1)$ .

$L^\infty(0, T; L^2)$  is not reflexive. But it is the dual space of  $L^1(0, T; L^2)$ , which is separable. So we can use Theorem 6.7 to find a sub-subsequence (a subsequence of the first subsequence) for which (relabelling) we also have  $u_{n_j} \overset{*}{\rightharpoonup} w$  in  $L^\infty(0, T; L^2)$ .  $L^2(0, T; L^2)$  is a Hilbert subspace of both these two spaces, so  $u_{n_j} \rightharpoonup u$  and  $u_{n_j} \overset{*}{\rightharpoonup} w$  in  $L^2(0, T; L^2)$ ; in a Hilbert space (in fact in any reflexive space) weak and weak-\* convergence are equivalent, so by uniqueness of weak limits  $u = w$ .

The space  $L^2(0, T; H^{-1})$  is the dual of  $L^2(0, T; H_0^1)$ , so we can use Theorem 6.7 again to find a weakly-\* convergent subsequence (we take a further sub-sub-sub-sequence so as not to lose the two ‘convergences’ that we had already).

What is, for example, weak-\* convergence in  $L^2(0, T; H_0^1)$ ? This is the dual space of  $L^2(0, T; H^{-1})$ ; so given any  $\varphi \in L^2(0, T; H^{-1})$ ,

$$\int_0^T \langle \dot{u}_{n_j}(t), \varphi(t) \rangle dt \rightarrow \int_0^T \langle v(t), \varphi(t) \rangle dt.$$

Similarly, weak convergence of  $u_{n_j}$  to  $u$  in  $L^2(0, T; H_0^1)$  means that for any  $\varphi \in L^2(0, T; H^{-1})$ ,

$$\int_0^T \langle \varphi(t), u_{n_j}(t) \rangle dt \rightarrow \int_0^T \langle \varphi(t), u(t) \rangle dt.$$

This is already enough to sort out one of our problems, whether  $v = \dot{u}$ . Take  $v \in H_0^1$  and  $\psi(t) \in C_c^\infty(0, T)$ , then, since  $v\psi(t) \in L^2(0, T; H_0^1)$ ,

$$\int_0^T \langle \dot{u}_{n_j}(t), v\psi(t) \rangle dt \rightarrow \int_0^T \langle v(t), v\psi(t) \rangle dt.$$

But

$$\begin{aligned} \int_0^T \langle \dot{u}_{n_j}(t), v\psi(t) \rangle dt &= \left\langle \int_0^T \dot{u}_{n_j}(t)\psi(t) dt, v \right\rangle \\ &= - \left\langle \int_0^T u_{n_j}(t)\dot{\psi}(t) dt, v \right\rangle = - \int_0^T \langle u_{n_j}(t), \dot{\psi}(t)v \rangle dt. \end{aligned}$$

Now certainly  $\psi(t)v \in L^2(0, T; H^{-1})$ , so using the weak convergence of  $u_{n_j}$  to  $u$  in  $L^2(0, T; H_0^1)$ , this expression converges to

$$- \int_0^T \langle u(t), \psi(t)v \rangle dt.$$

We therefore have, for every  $v \in H_0^1$ ,

$$\int_0^T \langle v(t), v\psi(t) \rangle dt = - \int_0^T \langle u(t), \psi(t)v \rangle dt.$$

This implies that

$$\left\langle \int_0^T v(t)\psi(t) dt, v \right\rangle = \left\langle - \int_0^T u(t)\dot{\psi}(t) dt, v \right\rangle$$

for every  $v \in H_0^1$ , i.e. that

$$\int_0^T v(t)\psi(t) dt = - \int_0^T u(t)\dot{\psi}(t) dt$$

as an equality in  $H^{-1}$ , so  $v = \dot{u}$ .

We want to show that every term in

$$\dot{u}_n - \Delta u_n = P_n f(t)$$

converges weakly-\* to its limit in  $L^2(0, T; H^{-1})$ . We have, at least, that  $\dot{u}_n \rightharpoonup v$  in this sense; and we showed before the  $P_n f(t) \rightarrow f(t)$  strongly in  $L^2(0, T; H^{-1})$ , and hence weakly. For the Laplacian term, take  $\varphi \in L^2(0, T; H_0^1)$ ; then

$$\begin{aligned} \int_0^T \langle -\Delta u_n(t), \varphi(t) \rangle dt &= \int_0^T \langle u_n(t), -\Delta \varphi(t) \rangle dt \\ &\rightarrow \int_0^T \langle u(t), -\Delta \varphi(t) \rangle dt \\ &= \int_0^T \langle -\Delta u(t), \varphi(t) \rangle dt, \end{aligned}$$

since  $\Delta \varphi(t) \in L^2(0, T; H^{-1})$  and  $u_n \rightharpoonup u$  in  $L^2(0, T; H_0^1)$ .

It follows that

$$\dot{u} - \Delta u = f(t)$$

as an equality in  $L^2(0, T; H^{-1})$ , with  $u \in L^2(0, T; H_0^1)$  and  $\dot{u} \in L^2(0, T; H^{-1})$ . We therefore have  $u \in C^0([0, T]; L^2)$ , and it only remains to check that  $u(0) = u_0$ .

Last time we used the uniform convergence of  $u_n$  to  $u$  (i.e. convergence in  $C^0([0, T]; L^2)$ ), but we need a more roundabout argument now.

First, since we have

$$\dot{u} - \Delta u = f(t)$$

in  $L^2(0, T; H^{-1})$ , choose  $v \in H_0^1$ , and then  $z \in C^1([0, T]; H_0^1) \subset L^2(0, T; H_0^1)$  with  $z(T) = 0$  and  $z(0) = v$ . Then

$$\int_0^T \langle \dot{u}(t), z(t) \rangle + (\nabla u(t), \nabla z(t) - \langle f(t), z(t) \rangle) dt = 0.$$

Integrating the first term by parts (since  $\dot{u}$  is the weak time derivative) we have

$$\int_0^T \langle u(t), -\dot{z}(t) \rangle + (\nabla u(t), \nabla z(t) - \langle f(t), z(t) \rangle) dt = (u(0), v).$$

(We are not just using the definition of the weak time derivative, but also the result of Lemma 5.2; see the examples.)

For the Galerkin approximations we have

$$\dot{u}_n - \Delta u_n = P_n f(t);$$

taking the inner product with  $z$  and integrating in space and time yields

$$\int_0^T \langle u_n(t), -\dot{z}(t) \rangle + (\nabla u_n(t), \nabla z(t)) - \langle P_n f(t), z(t) \rangle dt = (P_n u_0, v).$$

Using the fact that  $u_n \rightharpoonup u$  in  $L^2(0, T; H_0^1)$  and  $P_n f \rightarrow f$  in  $L^2(0, T; H^{-1})$ , we let  $n \rightarrow \infty$  to obtain

$$\int_0^T \langle u(t), -\dot{z}(t) \rangle + (\nabla u(t), \nabla z(t) - \langle f(t), z(t) \rangle) dt = (u_0, v).$$

This shows that  $(u(0), v) = (u_0, v)$  for every  $v \in H_0^1$ . So  $u(0) = u_0$  as required.