A Some Tools from Mathematical Analysis

In this appendix we give a short survey of results, essentially without proofs, from mathematical, particularly functional, analysis which are needed in our treatment of partial differential equations. We begin in Sect. A.1 with a simple account of abstract linear spaces with emphasis on Hilbert space, including the Riesz representation theorem and its generalization to bilinear forms of Lax and Milgram. We continue in Sect. A.2 with function spaces, where after a discussion of the spaces $\mathcal{C}^k$, integrability, and the L_p-spaces, we turn to L_2-based Sobolev spaces, with the trace theorem and Poincaré's inequality. The final Sect. A.3 is concerned with the Fourier transform.

A.1 Abstract Linear Spaces

Let V be a linear space (or vector space) with real scalars, i.e., a set such that if $u, v \in V$ and $\lambda, \mu \in \mathbf{R}$, then $\lambda u + \mu v \in V$. A *linear functional* (or *linear form*) L on V is a function $L : V \to \mathbf{R}$ such that

$$L(\lambda u + \mu v) = \lambda L(u) + \mu L(v), \quad \forall u, v \in V,\ \lambda, \mu \in \mathbf{R}.$$

A *bilinear form* $a(\cdot,\cdot)$ on V is a function $a : V \times V \to \mathbf{R}$, which is linear in each argument separately, i.e., such that, for all $u, v, w \in V$ and $\lambda, \mu \in \mathbf{R}$,

$$\begin{aligned} a(\lambda u + \mu v, w) &= \lambda a(u, w) + \mu a(v, w), \\ a(w, \lambda u + \mu v) &= \lambda a(w, u) + \mu a(w, v). \end{aligned}$$

The bilinear form $a(\cdot,\cdot)$ is said to be *symmetric* if

$$a(w, v) = a(v, w), \quad \forall v, w \in V,$$

and *positive definite* if

$$a(v, v) > 0, \quad \forall v \in V,\ v \neq 0.$$

A positive definite, symmetric, bilinear form on V is also called an *inner product* (or *scalar product*) on V. A linear space V with an inner product is called an *inner product space*.

If V is an inner product space and $(\cdot,\cdot)$ is an inner product on V, then we define the corresponding *norm* by

$$\|v\| = (v,v)^{1/2}, \quad \text{for } v \in V. \tag{A.1}$$

We recall the *Cauchy-Schwarz inequality*,

$$|(w,v)| \le \|w\|\,\|v\|, \quad \forall v,w \in V, \tag{A.2}$$

with equality if and only if $w = \lambda v$ or $v = \lambda w$ for some $\lambda \in \mathbf{R}$, and the *triangle inequality*,

$$\|w+v\| \le \|w\| + \|v\|, \quad \forall v,w \in V. \tag{A.3}$$

Two elements $v,w \in V$ for which $(v,w) = 0$ are said to be *orthogonal*.

An infinite sequence $\{v_i\}_{i=1}^\infty$ in V is said to converge to $v \in V$, also written $v_i \to v$ as $i \to \infty$ or $v = \lim_{i\to\infty} v_i$, if

$$\|v_i - v\| \to 0 \quad \text{as } i \to \infty.$$

The sequence $\{v_i\}_{i=1}^\infty$ is called a *Cauchy sequence* in V if

$$\|v_i - v_j\| \to 0 \quad \text{as } i,j \to \infty.$$

The inner product space V is said to be *complete* if every Cauchy sequence in V is convergent, i.e., if every Cauchy sequence $\{v_i\}_{i=1}^\infty$ has a limit $v = \lim v_i \in V$. A complete inner product space is called a *Hilbert space*.

When we want to emphasize that an inner product or a norm is associated to a specific space V, we write $(\cdot,\cdot)_V$ and $\|\cdot\|_V$.

It is sometimes important to permit the scalars in a linear space V to be complex numbers. Such a space is then an inner product space if there is a functional (v,w) defined on $V \times V$, which is linear in the first variable and hermitian, i.e., $(w,v) = \overline{(v,w)}$. The norm is then again defined by (A.1) and V is a complex Hilbert space if completeness holds with respect to this norm. For brevity we generally consider the case of real-valued scalars in the sequel.

More generally, a *norm* in a linear space V is a function $\|\cdot\| : V \to \mathbf{R}_+$ such that

$$\begin{aligned} \|v\| &> 0, && \forall v \in V,\ v \neq 0,\\ \|\lambda v\| &= |\lambda|\,\|v\|, && \forall \lambda \in \mathbf{R} \text{ (or } \mathbf{C}),\ v \in V,\\ \|v+w\| &\le \|v\| + \|w\|, && \forall v,w \in V. \end{aligned}$$

A function $|\cdot|$ is called a *seminorm* if these conditions hold with the exception that the first one is replaced by $|v| \ge 0,\ \forall v \in V$, i.e., if it is only positive semidefinite, and thus can vanish for some $v \neq 0$. A linear space with a norm is called a *normed linear space*. As we have seen, an inner product space is a normed linear space, but not all normed linear spaces are inner product spaces. A complete normed space is called a *Banach space*.

Let V be a Hilbert space and let $V_0 \subset V$ be a linear subspace. Such a subspace V_0 is said to be *closed* if it contains all limits of sequences in V_0, i.e., if $\{v_j\}_{j=1}^{\infty} \subset V_0$ and $v_j \to v$ as $j \to \infty$ implies $v \in V_0$. Such a V_0 is itself a Hilbert space, with the same inner product as V.

Let V_0 be a closed subspace of V. Then any $v \in V$ may be written uniquely as $v = v_0 + w$, where $v_0 \in V_0$ and w is orthogonal to V_0. The element v_0 may be characterized as the unique element in V_0 which is closest to v, i.e.,

$$\|v - v_0\| = \min_{u \in V_0} \|v - u\|. \tag{A.4}$$

This is called the *projection theorem* and is a basic result in Hilbert space theory. The element v_0 is called the *orthogonal projection* of v onto V_0 and is also denoted $P_{V_0}v$. One useful consequence of the projection theorem is that if the closed linear subspace V_0 is not equal to the whole space V, then it has a normal vector, i.e., there is a nonzero vector $w \in V$ which is orthogonal to V_0.

Two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are said to be *equivalent* in V if there are positive constants c and C such that

$$c\|v\|_b \leq \|v\|_a \leq C\|v\|_b, \quad \forall v \in V. \tag{A.5}$$

Let V, W be two Hilbert spaces. A linear operator $B : V \to W$ is said to be *bounded*, if there is a constant C such that

$$\|Bv\|_W \leq C\|v\|_V, \quad \forall v \in V. \tag{A.6}$$

The norm of a bounded linear operator B is

$$\|B\| = \sup_{v \in V \setminus \{0\}} \frac{\|Bv\|_W}{\|v\|_V}. \tag{A.7}$$

Thus

$$\|Bv\|_W \leq \|B\|\,\|v\|_V, \quad \forall v \in V,$$

and, by definition, $\|B\|$ is the smallest constant C such that (A.6) holds.

Note that a bounded linear operator $B : V \to W$ is continuous. In fact, if $v_j \to v$ in V, then $Bv_j \to Bv$ in W as $j \to \infty$, because

$$\|Bv_j - Bv\|_W = \|B(v_j - v)\|_W \leq \|B\|\,\|v_j - v\| \to 0, \quad \text{as } j \to \infty.$$

One can show that, conversely, a continuous linear operator is bounded.

In the special case that $W = \mathbf{R}$ the definition of an operator reduces to that of a linear functional. The set of all bounded linear functionals on V is called the *dual space* of V, denoted V^*. By (A.7) the norm in V^* is

$$\|L\|_{V^*} = \sup_{v \in V \setminus \{0\}} \frac{|L(v)|}{\|v\|_V}. \tag{A.8}$$

Note that V^* is itself a linear space if we define $(\lambda L + \mu M)(v) = \lambda L(v) + \mu M(v)$ for $L, M \in V^*$, $\lambda, \mu \in \mathbf{R}$. With the norm defined by (A.8), V^* is a normed linear space, and one may show that V^* is complete, and thus itself also a Banach space.

Similarly, we say that the bilinear form $a(\cdot,\cdot)$ on V is *bounded* if there is a constant M such that

$$(A.9) \qquad |a(w,v)| \le M\|w\|\,\|v\|, \quad \forall w, v \in V.$$

The next theorem states an important property of Hilbert spaces.

Theorem A.1. (Riesz' representation theorem.) *Let V be a Hilbert space with scalar product $(\cdot,\cdot)$. For each bounded linear functional L on V there is a unique $u \in V$ such that*

$$L(v) = (v,u), \quad \forall v \in V.$$

Moreover,

$$(A.10) \qquad \|L\|_{V^*} = \|u\|_V.$$

Proof. The uniqueness is clear since $(v,u_1) = (v,u_2)$ with $v = u_1 - u_2$ implies $\|u_1 - u_2\|^2 = (u_1 - u_2, u_1 - u_2) = 0$. If $L(v) = 0$ for all $v \in V$, then we may take $u = 0$. Assume now that $L(\bar v) \ne 0$ for some $\bar v \in V$. We will construct u as a suitably normalized "normal vector" to the "hyperplane" $V_0 = \{v \in V : L(v) = 0\}$, which is easily seen to be a closed subspace of V, see Problem A.2. Then $\bar v = v_0 + w$ with $v_0 \in V_0$ and w orthogonal to V_0 and $L(w) = L(\bar v) \ne 0$. But then $L(v - w\,L(v)/L(w)) = 0$, so that $(v - w\,L(v)/L(w), w) = 0$ and hence $L(v) = (v,u)$, $\forall v \in V$, where $u = w\,L(w)/\|w\|^2$. □

This result makes it natural to identify the linear functionals $L \in V^*$ with the associated $u \in V$, and thus V^* is equivalent to V, in the case of a Hilbert space.

We sometimes want to solve equations of the form: Find $u \in V$ such that

$$(A.11) \qquad a(u,v) = L(v), \quad \forall v \in V,$$

where V is a Hilbert space, L is a bounded linear functional on V, and $a(\cdot,\cdot)$ is a symmetric bilinear form, which is *coercive* in V, i.e.,

$$(A.12) \qquad a(v,v) \ge \alpha\|v\|_V^2, \quad \forall v \in V, \quad \text{with } \alpha > 0.$$

This implies that $a(\cdot,\cdot)$ is symmetric, positive definite, i.e., an inner product on V, and the Riesz representation theorem immediately gives the existence of a unique solution $u \in V$ for each $L \in V^*$.

Moreover, by taking $v = u$ in (A.11) we get

$$\alpha\|u\|_V^2 \le a(u,u) = L(u) \le \|L\|_{V^*}\,\|u\|_V,$$

so that, after cancelling one factor $\|u\|_V$,

$$\|u\|_V \le C\|L\|_{V^*}, \quad \text{where } C = 1/\alpha. \tag{A.13}$$

This is an example of an *energy estimate.*

If $a(\cdot,\cdot)$ is a symmetric bilinear form, which is coercive and bounded in V, so that (A.12) and (A.9) hold, then we may define a norm $\|\cdot\|_a$, the *energy norm*, by

$$\|v\|_a = a(v,v)^{1/2}, \quad \text{for } v \in V,$$

By (A.12) and (A.9) we then have

$$\sqrt{\alpha}\|v\|_V \le \|v\|_a \le \sqrt{M}\|v\|_V, \quad \forall v \in V, \tag{A.14}$$

and thus the norm $\|\cdot\|_a$ on V is equivalent to $\|\cdot\|_V$. Clearly, V is then also a Hilbert space with respect to the scalar product $a(\cdot,\cdot)$ and norm $\|\cdot\|_a$.

The solution of (A.11) may also be characterized in terms of a minimization problem.

Theorem A.2. *Assume that $a(\cdot,\cdot)$ is a symmetric, positive definite bilinear form and that L is a bounded linear form on the Hilbert space V. Then $u \in V$ satisfies* (A.11) *if and only if*

$$F(u) \le F(v), \quad \forall v \in V, \quad \text{where } F(v) = \tfrac{1}{2}a(v,v) - L(v). \tag{A.15}$$

Proof. Suppose first that u satisfies (A.11). Let $v \in V$ be arbitrary and define $w = v - u \in V$. Then $v = u + w$ and

$$\begin{aligned} F(v) &= \tfrac{1}{2}a(u+w,u+w) - L(u+w) \\ &= \tfrac{1}{2}a(u,u) - L(u) + a(u,w) - L(w) + \tfrac{1}{2}a(w,w) \\ &= F(u) + \tfrac{1}{2}a(w,w), \end{aligned}$$

where we have used (A.11) and the symmetry of $a(\cdot,\cdot)$. Since a is positive definite, this proves (A.15).

Conversely, if (A.15) holds, then for $v \in V$ given we have

$$g(t) := F(u+tv) \ge F(u) = g(0), \quad \forall\, t \in \mathbf{R},$$

so that $g(t)$ has a minimum at $t = 0$. But $g(t)$ is the quadratic polynomial

$$\begin{aligned} g(t) &= \tfrac{1}{2}a(u+tv,u+tv) - L(u+tv) \\ &= \tfrac{1}{2}a(u,u) - L(u) + t\big(a(u,v) - L(v)\big) + \tfrac{1}{2}t^2 a(v,v), \end{aligned}$$

and thus $0 = g'(0) = a(u,v) - L(v)$, which is (A.11). □

Thus, $u \in V$ satisfies (A.11) if and only if u minimizes the energy functional F. This method of studying the minimization problem by varying the argument of the functional F around the given vector u is called a variational method, and the equation (A.11) is called the *variational equation* of F.

The following theorem, which is known as the *Lax-Milgram lemma*, extends the Riesz representation theorem to nonsymmetric bilinear forms.

Theorem A.3. *If the bilinear form $a(\cdot,\cdot)$ is bounded and coercive in the Hilbert space V, and L is a bounded linear form in V, then there exists a unique vector $u \in V$ such that* (A.11) *is satisfied. Moreover, the energy estimate* (A.13) *holds.*

Proof. With $(\cdot,\cdot)$ the inner product in V we have by Riesz' representation theorem that there exists a unique $b \in V$ such that

$$L(v) = (b, v), \quad \forall v \in V.$$

Moreover, for each $u \in V$, $a(u,\cdot)$ is clearly also a bounded linear functional on V, so that there exists a unique $A(u) \in V$ such that

$$a(u, v) = (A(u), v), \quad \forall v \in V.$$

It is easy to check that $A(u)$ depends linearly and boundedly on u, so that $Au = A(u)$ defines $A : V \to V$ as a bounded linear operator. The equation (A.11) is therefore equivalent to $Au = b$, and to complete the proof of the theorem we shall show that this equation has a unique solution $u = A^{-1}b$ for each b.

Using the coercivity we have

$$\alpha \|v\|_V^2 \le a(v, v) = (Av, v) \le \|Av\|_V \|v\|_V,$$

so that

$$\alpha \|v\|_V \le \|Av\|_V, \quad \forall v \in V. \tag{A.16}$$

This shows uniqueness, since $Av = 0$ implies $v = 0$. This may also be expressed by saying that the null space $N(A) = \{v \in V : Av = 0\} = 0$, or that A is *injective*.

To show that there exists a solution u for each $b \in V$ means to show that each $b \in V$ belongs to the range $R(A) = \{w \in V : w = Av \text{ for some } v \in V\}$, i.e., $R(A) = V$, or A is *surjective*. To see this we first note that $R(A)$ is a closed linear subspace of V. To show that $R(A)$ is closed, assume that $Av_j \to w$ in V as $j \to \infty$. Then by (A.16) we have $\|v_j - v_i\|_V \le \alpha^{-1}\|Av_j - Av_i\|_V \to 0$ as $i, j \to \infty$. Hence $v_j \to v \in V$ as $j \to \infty$, and by the continuity of A, also $Av_j \to Av = w$. Therefore, $w \in R(A)$ and $R(A)$ is closed.

Assume now that $R(A) \ne V$. Then, by the projection theorem, there exists $w \ne 0$, which is orthogonal to $R(A)$. But, by the orthogonality,

$$\alpha \|w\|_V^2 \le a(w, w) = (Aw, w) = 0,$$

so that $w = 0$, which is a contradiction. Hence $R(A) = V$. This completes the proof that there is a unique solution for each $b \in V$. The energy estimate is proved in the same way as before. □

In the unsymmetric case there is no characterization of the solution in terms of energy minimization.

We finally make a remark about linear equations in finite-dimensional spaces. Let $V = \mathbf{R}^N$ and consider a linear equation in V, which may be written in matrix form as

$$Au = b,$$

where A is a $N \times N$ matrix and u, b are N-vectors. It is well-known that this equation has a unique solution $u = A^{-1}b$ for each $b \in V$, if the matrix A is nonsingular, i.e., if its determinant $\det(A) \neq 0$. If $\det(A) = 0$, then the homogeneous equation $Au = 0$ has nontrivial solutions $u \neq 0$, and $R(A) \neq V$ so that the inhomogeneous equation is not always solvable. Thus we have neither uniqueness nor existence for all $b \in V$. In particular, uniqueness only holds when $\det(A) \neq 0$, and we then also have existence. It is sometimes easy to prove uniqueness, and we then also obtain the existence of the solution at the same time.

A.2 Function Spaces

The Spaces $\mathcal{C}^k$

For $M \subset \mathbf{R}^d$ we denote by $\mathcal{C}(M)$ the linear space of continuous functions on M. The subspace $\mathcal{C}_\mathrm{b}(M)$ of all bounded functions is made into a normed linear space by setting (with a slight abuse of notation)

$$\|v\|_{\mathcal{C}(M)} = \sup_{x \in M} |v(x)|. \tag{A.17}$$

For example, this defines $\|v\|_{\mathcal{C}(\mathbf{R}^d)}$, which we use frequently. When M is a bounded and closed set, i.e., a compact set, the supremum in (A.17) is attained in M and we may write

$$\|v\|_{\mathcal{C}(M)} = \max_{x \in M} |v(x)|.$$

The norm (A.17) is therefore called the *maximum-norm*. Note that convergence in $\mathcal{C}(M)$,

$$\|v_i - v\|_{\mathcal{C}(M)} = \sup_{x \in M} |v_i(x) - v(x)| \to 0, \quad \text{as } i \to \infty,$$

is the same as uniform convergence in M. Recall that if a sequence of continuous functions is uniformly convergent in M, then the limit function is continuous. Using this fact it is not difficult to prove that $\mathcal{C}(M)$ is a complete normed space, i.e., a Banach space. $\mathcal{C}(M)$ is not a Hilbert space, because the maximum-norm is not associated with a scalar product as in (A.1).

Let now $\Omega \subset \mathbf{R}^d$ be a *domain*, i.e., a connected open set. For any integer $k \geq 0$, we denote by $\mathcal{C}^k(\Omega)$ the linear space of all functions v that are k times continuously differentiable in Ω, and by $\mathcal{C}^k(\bar{\Omega})$ the functions in $\mathcal{C}^k(\Omega)$, for which $D^\alpha v \in \mathcal{C}(\bar{\Omega})$ for all $|\alpha| \leq k$, where $D^\alpha v$ denotes the partial derivative of v defined in (1.8). If Ω is bounded, then the latter space is a Banach space with respect to the norm

$$\|v\|_{\mathcal{C}^k(\bar{\Omega})} = \max_{|\alpha| \leq k} \|D^\alpha v\|_{\mathcal{C}(\bar{\Omega})}.$$

For functions in $\mathcal{C}^k(\bar{\Omega})$, $k \geq 1$, we sometimes also use the seminorm containing only the derivatives of highest order,

$$|v|_{\mathcal{C}^k(\bar{\Omega})} = \max_{|\alpha| = k} \|D^\alpha v\|_{\mathcal{C}(\bar{\Omega})}.$$

A function has compact support in Ω if it vanishes outside some compact subset of Ω. We write $\mathcal{C}_0^k(\Omega)$ for the space of functions in $\mathcal{C}^k(\Omega)$ with compact support in Ω. In particular, such functions vanish near the boundary Γ, and for very large x if Ω is unbounded.

We say that a function is *smooth* if, depending on the context, it has sufficiently many continuous derivatives for the purpose at hand.

When there is no risk for confusion, we omit the domain of the functions from the notation of the spaces and write, e.g., $\mathcal{C}$ for $\mathcal{C}(\bar{\Omega})$ and $\|\cdot\|_{\mathcal{C}^k}$ for $\|\cdot\|_{\mathcal{C}^k(\bar{\Omega})}$, and similarly for other spaces that we introduce below.

Integrability, the Spaces L_p

Let Ω be a domain in $\mathbf{R}^d$. We shall need to work with integrals of functions $v = v(x)$ in Ω which are more general than those in $\mathcal{C}(\bar{\Omega})$. For a nonnegative function one may define the so-called *Lebesgue integral*

$$I_\Omega(v) = \int_\Omega v(x)\,\mathrm{d}x,$$

which may be either finite or infinite, and which agrees with the standard Riemann integral for $v \in \mathcal{C}(\bar{\Omega})$. The functions we consider are assumed measurable; we shall not go into details about this concept but just note that all functions that we encounter in this text will satisfy this requirement. A nonnegative function v is said to be integrable if $I_\Omega(v) < \infty$, and a general real or complex-valued function v is similarly integrable if $|v|$ is integrable. A subset Ω_0 of Ω is said to be a nullset, or a set of measure 0, if its volume $|\Omega|$ equals 0. Two functions which are equal except on a nullset are said to be equal almost everywhere (a.e.), and they then have the same integral. Thus if $v_1(x) = 1$ in a bounded domain Ω and if $v_2(x) = 1$ in Ω except at $x_0 \in \Omega$ where $v_2(x_0) = 2$, then $I_\Omega(v_1) = I_\Omega(v_2) = |\Omega|$. In particular, from the fact that a function is integrable we cannot draw any conclusion about its value at a point $x_0 \in \Omega$, i.e., the point values are not well defined. Also, since the boundary Γ of Ω is a nullset, $I_{\bar{\Omega}}(v) = I_\Omega(v)$ for any v.

We now define

$$\|v\|_{L_p} = \|v\|_{L_p(\Omega)} = \begin{cases} \Big(\int_\Omega |v(x)|^p \, dx\Big)^{1/p}, & \text{for } 1 \le p < \infty, \\ \operatorname{ess\,sup}_\Omega |v(x)|, & \text{for } p = \infty, \end{cases}$$

and say that $v \in L_p = L_p(\Omega)$ if $\|v\|_{L_p} < \infty$. Here the ess sup means the *essential supremum*, disregarding values on nullsets, so that, e.g., $\|v_2\|_{L_\infty} = 1$ for the function v_2 above, even though $\sup_\Omega v_2 = 2$. One may show that L_p is a complete normed space, i.e., a Banach space; the triangle inequality in L_p is called Minkowski's inequality. Clearly, any $v \in \mathcal{C}$ belongs to L_p for $1 \le p \le \infty$ if Ω is bounded, and

$$\|v\|_{L_p} \le C\|v\|_\mathcal{C}, \quad \text{with } C = |\Omega|^{1/p}, \quad \text{for } 1 \le p < \infty, \quad \text{and} \quad \|v\|_{L_\infty} = \|v\|_\mathcal{C},$$

but L_p also contains functions that are not continuous. Moreover, it is not difficult to show that $\mathcal{C}(\bar{\Omega})$ is incomplete with respect to the L_p-norm for $1 \le p < \infty$. To see this one constructs a sequence $\{v_i\}_{i=1}^\infty \subset \mathcal{C}(\bar{\Omega})$, which is a Cauchy sequence with respect to the L_p-norm, i.e., such that $\|v_i - v_j\|_{L_p} \to 0$, but whose limit $v = \lim_{i\to\infty} v_i$ is discontinuous. However, $\mathcal{C}(\bar{\Omega})$ is a *dense subspace* of $L_p(\Omega)$ for $1 \le p < \infty$, if Γ is sufficiently smooth. By this we mean that for any $v \in L_p$ there is a sequence $\{v_i\}_{i=1}^\infty \subset \mathcal{C}$ such that $\|v_i - v\|_{L_p} \to 0$ as $i \to \infty$. In other words, any function $v \in L_p$ can be approximated arbitrarily well in the L_p-norm by functions in $\mathcal{C}$ (in fact, for any k by functions in $\mathcal{C}_0^k$). In contrast, $\mathcal{C}$ is not dense in L_∞ since a discontinuous function cannot be well approximated uniformly by a continuous function.

The case L_2 is of particular interest to us, and this space is an inner product space, and hence a Hilbert space, with respect to the inner product

$$(v, w) = \int_\Omega v(x)w(x) \, dx. \tag{A.18}$$

In the case of complex-valued functions one takes the complex conjugate of $w(x)$ in the integrand.

Sobolev Spaces

We shall now introduce some particular Hilbert spaces which are natural to use in the study of partial differential equations. These spaces consist of functions which are square integrable together with their partial derivatives up to a certain order. To define them we first need to generalize the concept of a partial derivative.

Let Ω be a domain in $\mathbf{R}^d$ and let first $v \in \mathcal{C}^1(\bar{\Omega})$. Integration by parts yields

$$\int_\Omega \frac{\partial v}{\partial x_i}\phi\,\mathrm{d}x = -\int_\Omega v\frac{\partial \phi}{\partial x_i}\,\mathrm{d}x, \quad \forall \phi \in \mathcal{C}_0^1 = \mathcal{C}_0^1(\Omega).$$

If $v \in L_2 = L_2(\Omega)$, then $\partial v/\partial x_i$ does not necessarily exist in the classical sense, but we may define $\partial v/\partial x_i$ to be the linear functional

$$L(\phi) = \frac{\partial v}{\partial x_i}(\phi) = -\int_\Omega v\frac{\partial \phi}{\partial x_i}\,\mathrm{d}x, \quad \forall \phi \in \mathcal{C}_0^1. \tag{A.19}$$

This functional is said to be a *generalized* or *weak derivative* of v. When L is bounded in L_2, i.e., $|L(\phi)| \le C\|\phi\|$, it follows from Riesz' representation theorem that there exists a unique function $w \in L_2$, such that $L(\phi) = (w, \phi)$ for all $\phi \in L_2$, and in particular

$$-\int_\Omega v\frac{\partial \phi}{\partial x_i}\,\mathrm{d}x = \int_\Omega w\phi\,\mathrm{d}x, \quad \forall \phi \in \mathcal{C}_0^1.$$

We then say that the weak derivative belongs to L_2 and write $\partial v/\partial x_i = w$. In this case we thus have

$$\int_\Omega \frac{\partial v}{\partial x_i}\phi\,\mathrm{d}x = -\int_\Omega v\frac{\partial \phi}{\partial x_i}\,\mathrm{d}x, \quad \forall \phi \in \mathcal{C}_0^1. \tag{A.20}$$

In particular, if $v \in \mathcal{C}^1(\bar{\Omega})$, then the generalized derivative $\partial v/\partial x_i$ coincides with the classical derivative $\partial v/\partial x_i$.

In a similar way, with $D^\alpha v$ denoting the partial derivative of v defined in (1.8), we define the weak partial derivative $D^\alpha v$ as the linear functional

$$D^\alpha v(\phi) = (-1)^{|\alpha|}\int_\Omega v\, D^\alpha\phi\,\mathrm{d}x, \quad \forall \phi \in \mathcal{C}_0^{|\alpha|}. \tag{A.21}$$

When this functional is bounded in L_2, Riesz' representation theorem shows that there exists a unique function in L_2, which we denote by $D^\alpha v$, such that

$$(D^\alpha v, \phi) = (-1)^{|\alpha|}(v, D^\alpha \phi), \quad \forall \phi \in \mathcal{C}_0^{|\alpha|}.$$

We refer to Problem A.9 for further discussion of generalized functions.

We now define $H^k = H^k(\Omega)$, for $k \ge 0$, to be the space of all functions whose weak partial derivatives of order $\le k$ belong to L_2, i.e.,

$$H^k = H^k(\Omega) = \big\{v \in L_2 : D^\alpha v \in L_2 \text{ for } |\alpha| \le k\big\},$$

and we equip this space with the inner product

$$(v, w)_k = (v, w)_{H^k} = \sum_{|\alpha|\le k}\int_\Omega D^\alpha v D^\alpha w\,\mathrm{d}x,$$

and the corresponding norm

$$\|v\|_k = \|v\|_{H^k} = (v,v)_{H^k}^{1/2} = \Big(\sum_{|\alpha|\le k}\int_\Omega (D^\alpha v)^2\,\mathrm{d}x\Big)^{1/2}.$$

In particular, $\|v\|_0 = \|v\|_{L_2}$, and in this case we normally omit the subscript 0 and write $\|v\|$. Also

$$\|v\|_1 = \Big(\int_\Omega \Big\{v^2 + \sum_{j=1}^d \big(\frac{\partial v}{\partial x_j}\big)^2\Big\}\,\mathrm{d}x\Big)^{1/2} = \Big(\|v\|^2 + \|\nabla v\|^2\Big)^{1/2}$$

and

$$\|v\|_2 = \Big(\int_\Omega \Big\{v^2 + \sum_{j=1}^d \big(\frac{\partial v}{\partial x_j}\big)^2 + \sum_{i,j=1}^d \big(\frac{\partial^2 v}{\partial x_i \partial x_j}\big)^2\Big\}\,\mathrm{d}x\Big)^{1/2}.$$

We sometimes also use the seminorm, for $k \ge 1$,

(A.22) $$|v|_k = |v|_{H^k} = \Big(\sum_{|\alpha|=k}\int_\Omega (D^\alpha v)^2\,\mathrm{d}x\Big)^{1/2}.$$

Note that the seminorm vanishes for constant functions. Using the fact that L_2 is complete, one may show that H^k is complete and thus a Hilbert space, see Problem A.4. The space H^k is an example of a more general class of function spaces called Sobolev spaces.

It may be shown that $\mathcal{C}^l = \mathcal{C}^l(\bar\Omega)$ is dense in $H^k = H^k(\Omega)$ for any $l \ge k$, if Γ is sufficiently smooth. This is useful because it allows us to obtain certain results for H^k by carrying out the proof for functions in $\mathcal{C}^k$, which may be technically easier, and then extend the result to all $v \in H^k$ by using the density, cf. the proof of Theorem A.4 below.

Similarly, we denote by $W_p^k = W_p^k(\Omega)$ the normed space defined by the norm

$$\|v\|_{W_p^k} = \Big(\int_\Omega \sum_{|\alpha|\le k} |D^\alpha v|^p\,\mathrm{d}x\Big)^{1/p}, \quad \text{for } 1 \le p < \infty,$$

with the obvious modification for $p = \infty$. This space is in fact complete and hence a Banach space. For $p = 2$ we have $W_2^k = H^k$. Again, for $v \in \mathcal{C}^k$ we have $\|v\|_{W_\infty^k} = \|v\|_{\mathcal{C}^k}$.

Trace Theorems

If $v \in \mathcal{C}(\bar\Omega)$, then $v(x)$ is well defined for $x \in \Gamma$, the boundary of Ω. The *trace* γv of such a v on Γ is the restriction of v to Γ, i.e.,

(A.23) $$(\gamma v)(x) = v(x), \quad \text{for } x \in \Gamma.$$

Recall that since Γ is a nullset, the trace of $v \in L_2(\Omega)$ is not well defined.

Suppose now that $v \in H^1(\Omega)$. Is it then possible to define v uniquely on Γ, i.e., to define its trace γv on Γ? (One may show that functions in

$H^1(\Omega)$ are not necessarily continuous, cf. Theorem A.5 and Problems A.6, A.7 below.) This question can be made more precise by asking if it possible to find a norm $\|\cdot\|_{(\Gamma)}$ for functions on Γ and some constant C that

$$\|\gamma v\|_{(\Gamma)} \le C\|v\|_1, \quad \forall v \in \mathcal{C}^1(\bar{\Omega}). \tag{A.24}$$

An inequality of this form is called a *trace inequality.* If (A.24) holds, then by a density argument (see below) it is possible to extend the domain of definition of the trace operator γ from $\mathcal{C}^1(\bar{\Omega})$ to $H^1(\Omega)$, and (A.24) will also hold for all $v \in H^1(\Omega)$. The function space to which γv will belong will be defined by the norm $\|\cdot\|_{(\Gamma)}$ in (A.24).

We remark that in the above discussion the boundary Γ could be replaced by some other subset of Ω of dimension smaller than d.

In order to proceed with the trace theorems, we first consider a one-dimensional case, with Γ corresponding to a single point.

Lemma A.1. *Let $\Omega = (0,1)$. Then there is a constant C such that*

$$|v(x)| \le C\|v\|_1, \quad \forall x \in \bar{\Omega},\ \forall v \in \mathcal{C}^1(\bar{\Omega}).$$

Proof. For $x, y \in \Omega$ we have $v(x) = v(y) + \int_y^x v'(s)\,\mathrm{d}s$, and hence by the Cauchy-Schwarz inequality

$$|v(x)| \le |v(y)| + \int_0^1 |v'(s)|\,\mathrm{d}s \le |v(y)| + \|v'\|.$$

Squaring both sides and integrating with respect to y, we obtain,

$$v(x)^2 \le 2\big(\|v\|^2 + \|v'\|^2\big) = 2\|v\|_1^2. \tag{A.25}$$

which shows the desired estimate. □

We now show a simple trace theorem. By $L_2(\Gamma)$ we denote the Hilbert space of all functions that are square integrable on Γ with norm

$$\|w\|_{L_2(\Gamma)} = \Big(\int_\Gamma w^2\,\mathrm{d}s\Big)^{1/2}.$$

Theorem A.4. (Trace theorem.) *Let Ω be a bounded domain in $\mathbf{R}^d$ $(d \ge 2)$ with smooth or polygonal boundary Γ. Then the trace operator $\gamma : \mathcal{C}^1(\bar{\Omega}) \to \mathcal{C}(\Gamma)$ may be extended to $\gamma : H^1(\Omega) \to L_2(\Gamma)$, which defines the trace $\gamma v \in L_2(\Gamma)$ for $v \in H^1(\Omega)$. Moreover, there is a constant $C = C(\Omega)$ such that*

$$\|\gamma v\|_{L_2(\Gamma)} \le C\|v\|_1, \quad \forall v \in H^1(\Omega). \tag{A.26}$$

Proof. We first show the trace inequality (A.26) for functions $v \in \mathcal{C}^1(\bar{\Omega})$. For simplicity we consider only the case when $\Omega = (0,1) \times (0,1)$, the unit square in $\mathbf{R}^2$. The proof in the general case is similar. For $x = (x_1, x_2) \in \Omega$ we have by (A.25)

$$v(0, x_2)^2 \leq 2\Big(\int_0^1 v(x_1, x_2)^2 \, \mathrm{d}x_1 + \int_0^1 \Big(\frac{\partial v}{\partial x_1}(x_1, x_2) \Big)^2 \, \mathrm{d}x_1 \Big),$$

and hence by integration with respect to x_2,

$$\int_0^1 v(0, x_2)^2 \, \mathrm{d}x_2 \leq 2\big(\|v\|^2 + \|\nabla v\|^2 \big) = 2\|v\|_1^2.$$

The analogous estimates for the remaining parts of Γ complete the proof of (A.26) for $v \in \mathcal{C}^1$.

Let now $v \in H^1(\Omega)$. Since $\mathcal{C}^1$ is dense in H^1 there is a sequence $\{v_i\}_{i=1}^\infty \subset \mathcal{C}^1$ such that $\|v - v_i\|_1 \to 0$. This sequence is then a Cauchy sequence in H^1, i.e., $\|v_i - v_j\|_1 \to 0$ as $i, j \to \infty$. Applying (A.26) to $v_i - v_j \in \mathcal{C}^1$, we find

$$\|\gamma v_i - \gamma v_j\|_{L_2(\Gamma)} \leq C\|v_i - v_j\|_1 \to 0, \quad \text{as } i, j \to \infty,$$

i.e., $\{\gamma v_i\}_{i=1}^\infty$ is a Cauchy sequence in $L_2(\Gamma)$, and thus there exists $w \in L_2(\Gamma)$ such that $\gamma v_i \to w$ in $L_2(\Gamma)$ as $i \to \infty$. We define $\gamma v = w$. It is easy to show that (A.26) then holds for $v \in H^1$. This extends γ to a bounded linear operator $\gamma : H^1(\Omega) \to L_2(\Gamma)$. Since $\mathcal{C}^1$ is dense in H^1, there is only one such extension (prove this!). In particular, γ is independent of the choice of the sequence $\{v_i\}$. □

The constant in Theorem A.4 depends on the size and shape of the domain Ω. It is sometimes important to have more detailed information about this dependence, and in Problem A.15 we assume that the shape is fixed (a square) and investigate the dependence of the constant on the size of Ω.

The following result, of a somewhat similar nature, is a special case of the well-known and important Sobolev inequality.

Theorem A.5. *Let Ω be a bounded domain in $\mathbf{R}^d$ with smooth or polygonal boundary and let $k > d/2$. Then $H^k(\Omega) \subset \mathcal{C}(\bar{\Omega})$, and there exists a constant $C = C(\Omega)$ such that*

(A.27) $$\|v\|_{\mathcal{C}} \leq C\|v\|_k, \quad \forall v \in H^k(\Omega).$$

In the same way as for the trace theorem it suffices to show (A.27) for smooth v, see Problem A.20. The particular case when $d = k = 1$ is given in Lemma A.1, and Problem A.13 considers the case $\Omega = (0,1) \times (0,1)$. The general case is more complicated. As shown in Problems A.6, A.7, a function in $H^1(\Omega)$ with $\Omega \subset \mathbf{R}^d$ is not necessarily continuous when $d \geq 2$.

If we apply Sobolev's inequality to derivatives of v, we get

(A.28) $$\|v\|_{\mathcal{C}^\ell} \leq C\|v\|_k, \quad \forall v \in H^k(\Omega), \text{ if } k > \ell + d/2,$$

and we may similarly conclude that $H^k(\Omega) \subset \mathcal{C}^\ell(\bar{\Omega})$ if $k > \ell + d/2$.

The Space $H_0^1(\Omega)$. Poincaré's Inequality

Theorem A.4 shows that the trace operator $\gamma : H^1(\Omega) \to L_2(\Gamma)$ is a bounded linear operator. This implies that its null space,

$$H_0^1(\Omega) = \big\{v \in H^1(\Omega) : \gamma v = 0\big\},$$

is a closed subspace of $H^1(\Omega)$, and hence a Hilbert space with the norm $\|\cdot\|_1$. It is the set of functions in H^1 that vanish on Γ in the sense of trace. We note that the seminorm $|v|_1 = \|\nabla v\|$ defined in (A.22) is in fact a norm on $H_0^1(\Omega)$, equivalent to $\|\cdot\|_1$, as follows from the following result.

Theorem A.6. (Poincaré's inequality.) *If Ω is a bounded domain in $\mathbf{R}^d$, then there exists a constant $C = C(\Omega)$ such that*

$$\|v\| \le C\|\nabla v\|, \quad \forall v \in H_0^1(\Omega). \tag{A.29}$$

Proof. As an example we show the result for $\Omega = (0,1) \times (0,1)$. The proof in the general case is similar.

Since $\mathcal{C}_0^1$ is dense in H_0^1, it suffices to show (A.29) for $v \in \mathcal{C}_0^1$. For such a v we write

$$v(x) = \int_0^{x_1} \frac{\partial v}{\partial x_1}(s, x_2)\,\mathrm{d}s, \quad \text{for } x = (x_1, x_2) \in \Omega,$$

and hence by the Cauchy-Schwarz inequality

$$|v(x)|^2 \le \int_0^1 \mathrm{d}s \int_0^1 \Big(\frac{\partial v}{\partial x_1}(s, x_2)\Big)^2\,\mathrm{d}s.$$

The result now follows by integration with respect to x_2 and x_1, with $C = 1$ in this case. □

The equivalence of the norms $|\cdot|_1$ and $\|\cdot\|_1$ on $H_0^1(\Omega)$ now follows from

$$\|\nabla v\|^2 \le \|v\|_1^2 = \|v\|^2 + \|\nabla v\|^2 \le (C+1)\|\nabla v\|^2, \quad \forall v \in H_0^1(\Omega).$$

The dual space of $H_0^1(\Omega)$ is denoted $H^{-1}(\Omega)$. Thus $H^{-1} = (H_0^1)^*$ is the space of all bounded linear functionals on H_0^1. The norm in H^{-1} is (cf. (A.8))

$$\|L\|_{(H_0^1)^*} = \|L\|_{-1} = \sup_{v \in H_0^1} \frac{|L(v)|}{|v|_1}. \tag{A.30}$$

A.3 The Fourier Transform

Let v be a real or complex function in $L_1(\mathbf{R}^d)$. We define its Fourier transform for $\xi = (\xi_1, \dots, \xi_d) \in \mathbf{R}^d$ by

$$\mathcal{F}v(\xi) = \hat{v}(\xi) = \int_{\mathbf{R}^d} v(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x, \quad \text{where } x\cdot\xi = \sum_{j=1}^{d} x_j\xi_j.$$

The inverse Fourier transform is

$$\mathcal{F}^{-1}v(x) = \check{v}(x) = (2\pi)^{-d}\int_{\mathbf{R}^d} v(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\xi = (2\pi)^{-d}\hat{v}(-x), \quad \text{for } x \in \mathbf{R}^d.$$

If v and $\hat{v}$ are both in $L_1(\mathbf{R}^d)$, then Fourier's inversion formula holds, i.e.,

$$\mathcal{F}^{-1}(\mathcal{F}\, v) = (\hat{v})\check{} = v.$$

The inner product in $L_2(\mathbf{R}^d)$ of two functions can be expressed in terms of their Fourier transforms according to Parseval's formula,

$$\text{(A.31)} \qquad \int_{\mathbf{R}^d} v(x)\overline{w(x)}\,\mathrm{d}x = (2\pi)^{-d}\int_{\mathbf{R}^d} \hat{v}(\xi)\overline{\hat{w}(\xi)}\,\mathrm{d}\xi,$$

or

$$(v,w) = (2\pi)^{-d}(\hat{v},\hat{w}), \quad \text{where } (v,w) = (v,w)_{L_2(\mathbf{R}^d)}.$$

In particular, we have for the corresponding norms

$$\text{(A.32)} \qquad \|v\| = (2\pi)^{-d/2}\|\hat{v}\|.$$

Let $D^\alpha v$ be a partial derivative of v as defined in (1.8). We then have, assuming that v and its derivatives are sufficiently small for $|x|$ large,

$$\mathcal{F}(D^\alpha v)(\xi) = (i\xi)^\alpha\hat{v}(\xi) = i^{|\alpha|}\xi^\alpha\hat{v}(\xi), \quad \text{where } \xi^\alpha = \xi_1^{\alpha_1}\cdots\xi_d^{\alpha_d}.$$

In fact, by integration by parts,

$$\int_{\mathbf{R}^d} D^\alpha v(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x = (-1)^{|\alpha|}\int_{\mathbf{R}^d} v(x)D^\alpha(\mathrm{e}^{-\mathrm{i}x\cdot\xi})\,\mathrm{d}x = (i\xi)^\alpha\hat{v}(\xi).$$

Further, translation of the argument of the function corresponds to multiplication of its Fourier transform by an exponential,

$$\text{(A.33)} \qquad \mathcal{F}v(\cdot + y)(\xi) = \mathrm{e}^{\mathrm{i}y\cdot\xi}\hat{v}(\xi), \quad \text{for } y \in \mathbf{R}^d,$$

and for scaling of the argument we have

$$\text{(A.34)} \qquad \mathcal{F}v(a\cdot)(\xi) = a^{-d}\hat{v}(a^{-1}\xi), \quad \text{for } a > 0.$$

The convolution of two functions v and w is defined by

$$(v * w)(x) = \int_{\mathbf{R}^d} v(x-y)w(y)\,\mathrm{d}y = \int_{\mathbf{R}^d} v(y)w(x-y)\,\mathrm{d}y,$$

and we have

$$\mathcal{F}(v * w)(\xi) = \hat{v}(\xi)\hat{w}(\xi),$$

because

$$\begin{aligned}\int_{\mathbf{R}^d} \Big(\int_{\mathbf{R}^d} & v(x-y)w(y)\,\mathrm{d}y\Big)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \\ &= \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} v(x-y)w(y)\mathrm{e}^{-\mathrm{i}(x-y)\cdot\xi}\mathrm{e}^{-\mathrm{i}y\cdot\xi}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} v(z)w(y)\mathrm{e}^{-\mathrm{i}z\cdot\xi}\mathrm{e}^{-\mathrm{i}y\cdot\xi}\,\mathrm{d}z\,\mathrm{d}y.\end{aligned}$$

It follows, which can also easily be shown directly, that differentiation of a convolution can be carried out on either factor,

$$D^\alpha(v * w) = D^\alpha v * w = v * D^\alpha w.$$

A.4 Problems

Problem A.1. Let V be a Hilbert space with scalar product $(\cdot,\cdot)$ and let $u \in V$ be given. Define $L : V \to \mathbf{R}$ by $L(v) = (u, v)\ \forall v \in V$. Prove that L is a bounded linear functional on V. Determine $\|L\|$.

Problem A.2. Prove that if $L : V \to \mathbf{R}$ is a bounded linear functional and $\{v_i\}$ is a sequence with $L(v_i) = 0$ that converges to $v \in V$, then $L(v) = 0$. This proves that the subspace V_0 in the proof of Theorem A.1 is closed.

Problem A.3. Prove the energy estimate (A.13) by using (A.10) and (A.14). Hint: Recall (A.8) and note that (A.10) means

$$\sup_{v \in V\setminus\{0\}} \frac{|L(v)|}{\|v\|_a} = \|u\|_a.$$

Problem A.4. Given that $L_2(\Omega)$ is complete, prove that $H^1(\Omega)$ is complete. Hint: Assume that $\|v_j - v_i\|_1 \to 0$ as $i, j \to \infty$. Show that there are v, w_k such that $\|v_j - v\| \to 0$, $\|\partial v_j/\partial x_k - w_k\| \to 0$, and that $w_k = \partial v/\partial x_k$ in the sense of weak derivative.

Problem A.5. Let $\Omega = (-1, 1)$ and let $v : \Omega \to \mathbf{R}$ be defined by $v(x) = 1$ if $x \in (-1, 0)$ and $v(x) = 0$ if $x \in (0, 1)$. Prove that $v \in L_2(\Omega)$ and that v can be approximated arbitrarily well in L_2-norm by $\mathcal{C}^0$-functions.

Problem A.6. Let Ω be the unit ball in $\mathbf{R}^d$, $d = 1, 2, 3$, i.e., $\Omega = \{x \in \mathbf{R}^d : |x| < 1\}$. For which values of $\lambda \in \mathbf{R}$ does the function $v(x) = |x|^\lambda$ belong to (a) $L_2(\Omega)$, (b) $H^1(\Omega)$?

Problem A.7. Check if the function $v(x) = \log(-\log|x|^2)$ belongs to $H^1(\Omega)$ if $\Omega = \{x \in \mathbf{R}^2 : |x| < \frac{1}{2}\}$. Are functions in $H^1(\Omega)$ necessarily bounded and continuous?

Problem A.8. It is known that $\mathcal{C}_0^1(\Omega)$ is dense in $L_2(\Omega)$ and $H_0^1(\Omega)$. Explain why $\mathcal{C}_0^1(\Omega)$ is not dense in $H^1(\Omega)$.

Problem A.9. The generalized (or weak) derivative defined in (A.19) is a special case of the so-called *generalized functions* or *distributions*. Another important example is *Dirac's delta*, which is defined as a linear functional acting on continuous test functions, for $\Omega \subset \mathbf{R}^d$,

$$\delta(\phi) = \phi(0), \quad \forall \phi \in \mathcal{C}_0^0(\Omega).$$

Let now $d = 1$, $\Omega = (-1, 1)$ and

$$f(x) = \begin{cases} x, & x \geq 0, \\ 0, & x \leq 0, \end{cases} \qquad g(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases}$$

Show that $f' = g$, $g' = \delta$ in the sense of generalized derivative, i.e.,

$$\begin{aligned} f'(\phi) &= -\int_\Omega f\phi' \,\mathrm{d}x = \int_\Omega g\phi \,\mathrm{d}x, \qquad && \forall \phi \in \mathcal{C}_0^1(\Omega), \\ g'(\phi) &= -\int_\Omega g\phi' \,\mathrm{d}x = \phi(0), && \forall \phi \in \mathcal{C}_0^1(\Omega). \end{aligned}$$

Conclude that the generalized derivative $f' = g$ belongs to L_2, but that $g' = \delta$ does not. For the latter statement, you must show that δ is not bounded with respect to the L_2-norm, i.e., you need to find a sequence of test functions such that $\|\phi_i\|_{L_2} \to 0$, but $\phi_i(0) = 1$ as $i \to \infty$. Thus, $f \in H^1(\Omega)$ and $g \notin H^1(\Omega)$.

Problem A.10. For $f \in L_2(\Omega)$ we define the linear functional $f(v) = (f, v)$ $\forall v \in L_2(\Omega)$. Show the inequality, cf. (A.30),

$$\|f\|_{-1} \leq C\|f\|, \quad \forall f \in L_2(\Omega).$$

Conclude that $L_2(\Omega) \subset H^{-1}(\Omega)$.

Problem A.11. Let $\Omega = (0, 1)$ and $f(x) = 1/x$. Show that $f \notin L_2(\Omega)$. Show that $f \in H^{-1}(\Omega)$ by defining the linear functional $f(v) = (f, v)$ $\forall v \in H_0^1(\Omega)$, and proving the inequality

$$|(f, v)| \leq C\|v'\|, \quad \forall v \in H_0^1(\Omega).$$

Conclude that $H^{-1}(\Omega) \not\subset L_2(\Omega)$.

Problem A.12. Prove that if $\Omega = (0, L)$ is a finite interval, then there is a constant $C = C(L)$ such that, for all $x \in \bar{\Omega}$ and $v \in \mathcal{C}^1(\bar{\Omega})$,

(a) $$|v(x)| \leq L^{-1} \int_\Omega |v| \,\mathrm{d}y + \int_\Omega |v'| \,\mathrm{d}y \leq C\|v\|_{W_1^1(\Omega)},$$

(b) $$|v(x)|^2 \leq L^{-1} \int_\Omega |v|^2 \,\mathrm{d}y + L\int_\Omega |v'|^2 \,\mathrm{d}y \leq C\|v\|_1^2,$$

(c) $$|v(x)|^2 \leq L^{-1}\|v\|^2 + 2\|v\|\,\|v'\| \leq C\|v\|\,\|v\|_1.$$

Problem A.13. Prove that if Ω is the unit square in $\mathbf{R}^2$, then there exists a constant C such that

(a) $\|v\|_{L_1(\Gamma)} \leq C\|v\|_{W_1^1(\Omega)}, \quad \forall v \in \mathcal{C}^1(\bar{\Omega}),$

(b) $\|v\|_{\mathcal{C}} \leq C\|v\|_{W_1^2}, \quad \forall v \in \mathcal{C}^2(\bar{\Omega}).$

Since $\|v\|_{W_1^2} \leq 3^{1/2}|\Omega|^{1/2}\|v\|_{H^2}$, part (b) implies the special case of Theorem A.5 with $k = d = 2$ and Ω a square domain. Part (b) directly generalizes to $\|v\|_{\mathcal{C}} \leq C\|v\|_{W_1^d}$ for $\Omega \subset \mathbf{R}^d$. Hint: Proof of Theorem A.4.

Problem A.14. (Scaling of Sobolev norms.) Let L be a positive number and consider the coordinate transformation $x = L\hat{x}$, which maps the bounded domain $\Omega \subset \mathbf{R}^d$ onto $\hat{\Omega}$. A function v defined on Ω is transformed to a function $\hat{v}$ on $\hat{\Omega}$ according to $\hat{v}(\hat{x}) = v(L\hat{x})$. Prove the scaling identities

(a) $\|v\|_{L_2(\Omega)} = L^{d/2}\|\hat{v}\|_{L_2(\hat{\Omega})},$

(b) $\|\nabla v\|_{L_2(\Omega)} = L^{d/2-1}\|\hat{\nabla}\hat{v}\|_{L_2(\hat{\Omega})},$

(c) $\|v\|_{L_2(\Gamma)} = L^{d/2-1/2}\|\hat{v}\|_{L_2(\hat{\Gamma})}.$

Problem A.15. (Scaled trace inequality.) Let $\Omega = (0, L) \times (0, L)$ be a square domain of side L. Prove the scaled trace inequality

$$\|v\|_{L_2(\Gamma)} \leq C\Big(L^{-1}\|v\|^2_{L_2(\Omega)} + L\|\nabla v\|^2_{L_2(\Omega)}\Big)^{1/2}, \quad \forall v \in \mathcal{C}^1(\bar{\Omega}).$$

Hint: Apply (A.26) with $\hat{\Omega} = (0, 1) \times (0, 1)$ and use the scaling identities in Problem A.14.

Problem A.16. Let Ω be the unit square in $\mathbf{R}^2$. Prove the trace inequality in the form

$$\|v\|^2_{L_2(\Gamma)} \leq C\big(\|v\|^2_{L_2(\Omega)} + \|v\|_{L_2(\Omega)}\|\nabla v\|_{L_2(\Omega)}\big) \leq C\|v\|\,\|v\|_1.$$

Hint: Start from

$$v(0, y_2)^2 = v(y_1, y_2)^2 - \int_0^{y_1} \frac{\partial}{\partial x_1} v(s, y_2)^2 \, ds.$$

Problem A.17. It is a fact from linear algebra that all norms on a finite-dimensional space V are equivalent. Illustrate this by proving the following norm equivalences in $V = \mathbf{R}^N$:

(A.35) $\|v\|_{l_2} \leq \|v\|_{l_1} \leq \sqrt{N}\|v\|_{l_2},$

(A.36) $\|v\|_{l_\infty} \leq \|v\|_{l_2} \leq \sqrt{N}\|v\|_{l_\infty},$

(A.37) $\|v\|_{l_\infty} \leq \|v\|_{l_1} \leq N\|v\|_{l_\infty},$

where

$$\|v\|_{l_p} = \Big(\sum_{j=1}^{N} |v_j|^p\Big)^{1/p} \quad \text{for } 1 \le p < \infty, \quad \|v\|_{l_\infty} = \max_{1\le j\le N} |v_j|.$$

Note that the equivalence constants tend to infinity as $N \to \infty$.

Problem A.18. Prove (A.33) and (A.34).

Problem A.19. Prove that the Fourier transform of $v(x) = \mathrm{e}^{-|x|^2}$ is $\hat{v}(\xi) = \pi^{d/2}\mathrm{e}^{-|\xi|^2/4}$.

Problem A.20. Assume that Sobolev's inequality in (A.27) has been proved for all $v \in \mathcal{C}^k(\bar{\Omega})$ with $k > d/2$. Prove Sobolev's imbedding $H^k(\Omega) \subset \mathcal{C}(\bar{\Omega})$. In other words, for each $v \in H^k(\Omega)$ show that there is $w \in \mathcal{C}(\bar{\Omega})$ such that $v = w$ almost everywhere, i.e., $\|v - w\|_{L_2} = 0$. Hint: $\mathcal{C}^k(\bar{\Omega})$ is dense in $H^k(\Omega)$ and $\mathcal{C}(\bar{\Omega})$ is a Banach space.

B Orientation on Numerical Linear Algebra

Both finite difference and finite element methods for elliptic problems lead to linear algebraic systems of equations of the form

$$AU = b, \tag{B.1}$$

where A is a nonsingular square matrix of order N. Also in time-stepping methods for evolution equations, problems of elliptic type need to be solved in the successive time steps. To solve such systems efficiently therefore becomes an important part of numerical analysis. When the dimension of the computational domain is at least 2 this is normally not possible by direct methods, and, except in special cases, one therefore turns to iterative methods. These take advantage of the fact that the matrices involved are sparse, i.e., most of their elements are zero, and have other special features. In this appendix we give a short overview, without proofs, of the most commonly used methods.

B.1 Direct Methods

We consider first the case that the system (B.1) derives from the standard finite difference approximation (4.3) of the two-point boundary value problem (4.1). In this case A is a tridiagonal matrix, and it is easy to see that A may then be factored in $O(N)$ algebraic operations as $A = LR$, where L is bidiagonal and lower triangular and R is bidiagonal, upper triangular. The system may thus be written

$$LRU = b,$$

and one may now first solve $LG = b$ for $G = RU$ in $O(N)$ operations and then solve the latter equation for U, also in $O(N)$ operations. Altogether this is a direct method for (B.1), which requires $O(N)$ operations. Since the number of unknowns is N, this is the smallest possible order for any method.

Consider now an elliptic problem in a domain $\Omega \subset \mathbf{R}^d$ with $d \geq 2$. Using either finite differences or finite elements based on a quasi-uniform family of meshes, the dimension N of the corresponding finite dimensional problem is of order $O(h^{-d})$, where h is the mesh-size, and for $d \geq 2$ direct solution by Gauss elimination is normally not feasible as this method requires

$O(N^3) = O(h^{-3d})$ algebraic operations. Except in special cases one therefore turns to iterative methods.

One case when a direct method can be used, however, is provided by the model problem (4.11) with the five-point finite difference scheme on the unit square, which may be solved directly by using the discrete Fourier transform, defined by

$$\hat{b}_m = \sum_j b_j \mathrm{e}^{-2\pi \mathrm{i}\, m\cdot j\, h}, \quad m = (m_1, m_2),\ j = (j_1, j_2).$$

In fact, we then have $(-\Delta_h U)\hat{}_m = 2\pi^2|m|^2\hat{U}_m$, hence $\hat{U}_m = (2\pi^2|m|^2)^{-1}\hat{b}_m$, so that by the inverse discrete Fourier transform

$$U^j = \sum_m (2\pi^2|m|^2)^{-1}\hat{b}_m \mathrm{e}^{2\pi \mathrm{i}\, m\cdot j\, h}.$$

Using the Fast Fourier Transform (FFT) both $\hat{b}_m$ and U^j may be calculated in $O(N \log N)$ operations.

B.2 Iterative Methods. Relaxation, Overrelaxation, and Acceleration

As a basic iterative method for (B.1) we consider the Richardson method

$$U^{n+1} = U^n - \tau(AU^n - b) \quad \text{for } n \geq 0, \quad \text{with } U^0 \text{ given}, \tag{B.2}$$

where τ is a positive parameter. With U the exact solution of (B.1) we have

$$U^n - U = R(U^{n-1} - U) = \cdots = R^n(U^0 - U), \quad \text{where } R = I - \tau A,$$

and hence the rate of convergence of the method depends on $\|R^n\|$, where $\|M\| = \max_{\|x\|=1}\|Mx\|$ is the matrix norm subordinate to the Euclidean norm $\|\cdot\|$ in $\mathbf{R}^N$. When A is symmetric positive definite (SPD) and has eigenvalues $\{\lambda_j\}_{j=1}^N$, then, since $\{1 - \tau\lambda_j\}_{j=1}^N$ are the eigenvalues of R, we have

$$\|R^n\| = \rho^n, \quad \text{where } \rho = \rho(R) = \max_i |1 - \tau\lambda_i|,$$

and (B.2) converges if $\rho < 1$. The choice of τ which gives the smallest value of ρ is $\tau = 2/(\lambda_1 + \lambda_N)$, in which case $\rho = (\kappa - 1)/(\kappa + 1)$, where $\kappa = \kappa(A) = \lambda_N/\lambda_1$ is the condition number or A. We note, however, that this choice of τ requires knowledge of λ_1 and λ_N which is not normally available. In applications to second order elliptic problems one often has $\kappa = O(h^{-2})$ so that $\rho \leq 1 - ch^2$ with $c > 0$. Hence with the optimal choice of τ the number of iterations required to reduce the error to a small $\epsilon > 0$ is of order $O(h^{-2}|\log\epsilon|)$. Since each iteration uses $O(h^{-d})$ operations in the application of $I - \tau A$, this shows that the total number of operations needed to reduce the error to a given tolerance is of order $O(h^{-d-2})$, which is smaller than for the direct solution by Gauss elimination when $d \geq 2$.

The early more refined methods were designed for finite difference methods of positive type for second order elliptic equations, particularly for the five-point scheme (4.12). The corresponding matrix may then be written $A = D - E - F$, where D is diagonal and E and F are (elementwise) nonnegative and strictly lower and upper triangular. Examples of more efficient methods are then the Jacobi and Gauss-Seidel methods which are defined by

$$U^{n+1} = U^n - B(AU^n - b) = RU^n + Bb, \quad \text{with } R = I - BA, \tag{B.3}$$

in which $B = B_{\mathrm{J}} = D^{-1}$ or $B = B_{\mathrm{GS}} = (D - E)^{-1}$, so that $R = R_{\mathrm{J}} = D^{-1}(E + F)$ and $R = R_{\mathrm{GS}} = (D - E)^{-1}F$, respectively. In the application to the model problem (4.9) in the unit square, using the five-point operator, the equations may be normalized so that $D = 4I$ and the application of R_{J} then simply means that the new value in the iteration step at any interior mesh-point x_j is obtained by taking the average of the old values at the four neighboring points $x_{j\pm e_l}$. The Gauss-Seidel method also takes averages, but with the mesh-points taken in a given order, and successively uses the values already obtained in forming the averages. The methods are therefore also referred to as the methods of simultaneous and successive displacements, respectively. For the model problem one may easily determine the eigenvalues and eigenvectors of A and show that with $h = 1/M$ one has $\rho(R_{\mathrm{J}}) = \cos(\pi h) = 1 - \frac{1}{2}\pi^2 h^2 + O(h^4)$ and $\rho(R_{\mathrm{GS}}) = \rho(R_{\mathrm{J}})^2 = 1 - \pi^2 h^2 + O(h^4)$, so that the number of iterations needed to reduce the error to ϵ is of the orders $2h^{-2}\pi^2|\log\epsilon|$ and $h^{-2}\pi^2|\log\epsilon|$, respectively. The Gauss-Seidel method thus requires about half as many iterations as the Jacobi method.

Forming the averages in the Jacobi and Gauss-Seidel methods may be thought as relaxation. It turns out that one may obtain better results than those described above by overrelaxation, i.e., by choosing

$$B_\omega = (D - \omega E)^{-1} \quad \text{and} \quad R_\omega = (D - \omega E)^{-1}((1-\omega)E + F), \quad \text{with } \omega > 1.$$

It may be shown that for the model problem the optimal choice of the parameter is

$$\omega_{\mathrm{opt}} = 2/(1 + \sqrt{1 - \rho^2}), \quad \text{where } \rho = \rho(B_{\mathrm{J}}) = \cos(\pi h),$$

i.e., $\omega_{\mathrm{opt}} = 2/(1 + \sin(\pi h)) = 2 - 2\pi h + O(h^2)$, and that correspondingly

$$\rho(R_{\omega_{\mathrm{opt}}}) = \omega_{\mathrm{opt}} - 1 = 1 - 2\pi h + O(h^2).$$

The number of iterations required is thus then of order $O(h^{-1})$, which is significantly smaller than for the above methods. This is the method of successive overrelaxation (SOR).

We consider again an iterative method of the form (B.3) with $\rho(R) < 1$. For the purpose of accelerating the convergence we now introduce the new sequence $V^n = \sum_{j=0}^n \beta_{nj} U^j$, where the β_{nj} are real numbers. Setting $p_n(\lambda) = \sum_{j=0}^n \beta_{nj} \lambda^j$, and assuming $p_n(1) = \sum_{j=0}^n \beta_{nj} = 1$ for $n \geq 0$, we obtain easily $V^n - U = p_n(R)(U^0 - U)$, where U is the solution of (B.1). For V^n to converge fast to U one therefore wants to choose the β_{nj} in such a way that the spectral radius $\rho(p_n(R))$ becomes small with n. By the Cayley-Hamilton theorem for matrices one has $p_N(R) = 0$, if p_N is the characteristic polynomial of R, and hence $V^N = U$, but this requires a prohibitively large number of iterations. For $n < N$ we have by the spectral mapping theorem that $\rho(p_n(R)) = \max_i |p_n(\mu_i)|$, where $\{\mu_i\}_{i=1}^N$ are the eigenvalues of R. In particular, if R is symmetric and $\rho = \rho(R)$, so that $|\mu_i| \leq \rho$ for all i, then one may show that, taking the maximum instead over $[-\rho, \rho] \supset \sigma(R)$, the optimal polynomial is $p_n(\lambda) = T_n(\lambda/\rho)/T_n(1/\rho)$, where T_n is the nth Chebyshev polynomial, and the corresponding value of $\rho(p_n(R))$ is bounded by

$$\begin{aligned} T_n(1/\rho)^{-1} &= 2\Big\{\Big(\frac{1+\sqrt{1-\rho^2}}{\rho}\Big)^n + \Big(\frac{1+\sqrt{1-\rho^2}}{\rho}\Big)^{-n}\Big\}^{-1} \\ &\leq 2\Big(\frac{\rho}{1+\sqrt{1-\rho^2}}\Big)^n. \end{aligned}$$

For the model problem using the Gauss-Seidel basic iteration we have as above $\rho = 1 - \pi^2 h^2 + O(h^4)$ and we find that the average error reduction factor per iteration step in our present method is bounded by $1 - \sqrt{2}\pi h + O(h^2)$, which is of the same order of magnitude as for SOR.

B.3 Alternating Direction Methods

We now describe the Peaceman-Rachford alternating direction implicit iterative method for the model problem (4.9) on the unit square, using the five-point discrete elliptic equation (4.11) with $h = 1/M$. In this case we may write $A = H + V$, where H and V correspond to the horizontal and vertical difference operators $-h^2\partial_1\bar{\partial}_1$ and $-h^2\partial_2\bar{\partial}_2$. Note that H and V are positive definite and commute. Introducing an acceleration parameter τ and an intermediate value $U^{n+1/2}$, we may consider the scheme defining U^{n+1} from U^n by

$$\begin{aligned} (\tau I + H)U^{n+1/2} &= (\tau I - V)U^n + b, \\ (\tau I + V)U^{n+1} &= (\tau I - H)U^{n+1/2} + b, \end{aligned} \tag{B.4}$$

or after elimination, with G_τ appropriate and using that H and V commute,

$$U^{n+1} = R_\tau U^n + G_\tau, \quad \text{where } R_\tau = (\tau I - H)(\tau I + H)^{-1}(\tau I - V)(\tau I + V)^{-1}.$$

Note that the equations in (B.4) have tridiagonal matrices and may be solved in $O(N)$ operations, as we have indicated earlier. The error satisfies $U^n - U = R_\tau^n(U^0 - U)$, and with μ_i the (common) eigenvalues of H and V, we have $\|R_\tau\| \le \max_i |(\tau - \mu_i)/(\tau + \mu_i)|^2 < 1$, where it is easy to see that the maximum occurs for $i = 1$ or M. With $\mu_1 = 4\sin^2(\frac{1}{2}\pi h)$, $\mu_M = 4\cos^2(\frac{1}{2}\pi h)$ the optimal τ is $\tau_{\text{opt}} = (\mu_1 \mu_M)^{1/2}$ with the maximum for $i = 1$, so that, with $\kappa = \kappa(H) = \kappa(V) = \mu_M/\mu_1$,

$$\|R_{\tau_{\text{opt}}}\| \le \Big(\frac{(\mu_1\mu_M)^{1/2} - \mu_1}{(\mu_1\mu_M)^{1/2} + \mu_1}\Big)^{1/2} = \frac{\kappa^{1/2} - 1}{\kappa^{1/2} + 1} = 1 - \pi h + O(h^2).$$

This again shows the same order of convergence as for SOR.

A more efficient procedure is obtained by using varying acceleration parameters τ_j, $j = 1, 2, \ldots$, corresponding to the n step error reduction matrix $\tilde{R}_n = \prod_{j=1}^n R_{\tau_j}$. It can be shown that the τ_j can be chosen cyclically with period m in such a way that $m \approx c \log \kappa \approx c \log(1/h)$, so that the average error reduction rate is

$$\|\tilde{R}_m\|^{1/m} = \max_{1 \le i \le M} \Big(\prod_{j=0}^{m-1} \Big|\frac{\tau_j - \mu_i}{\tau_j + \mu_i}\Big|\Big)^{2/m} \le 1 - c(\log(1/h))^{-1}, \quad c > 0.$$

The analysis indicated depends strongly on the fact that H and V commute, which only happens for rectangles and constant coefficients, but the method may be defined and shown convergent in more general cases.

B.4 Preconditioned Conjugate Gradient Methods

We now turn to some iterative methods for systems mainly associated with the emergence of the finite element method. We begin by describing the conjugate gradient method, and assume that A is SPD. Considering the iterative method for (B.1) defined by

$$U^{n+1} = (I - \tau_n A)U^n + \tau_n b \quad \text{for } n \ge 0, \quad \text{with } U^0 = 0,$$

we find at once that, for any choice of the parameters τ_j, U^n belongs to the so-called Krylov space $K_n(A; b) = \text{span}\{b, Ab, \ldots, A^{n-1}b\}$, i.e., consisting of linear combinations of the $A^i b$, $i = 0, \ldots, n-1$. The conjugate gradient method defines these parameters so that U^n is the best approximation of the exact solution U of (B.1) in $K_n(A; b)$ with respect to the norm defined by $|U| = (AU, U)^{1/2}$, i.e., U^n is the orthogonal projection of U onto $K_n(A; b)$ with respect to the inner product (AV, W). By our above discussion it follows that, with $\kappa = \kappa(A)$ the condition number of A,

$$\text{(B.5)} \qquad |U^n - U| \le (T_n(1/\rho))^{-1}|U| \le 2\Big(\frac{\kappa^{1/2} - 1}{\kappa^{1/2} + 1}\Big)^n |U|.$$

The computation of U^n can be carried out by a two term recurrence relation, for instance, in the following form using the residuals $r^n = b - AU^n$ and the auxiliary vectors $q^n \in K_{n+1}(A;b)$, orthogonal to $K_n(A;b)$,

$$U^{n+1} = U^n + \frac{(r^n, q^n)}{(Aq^n, q^n)} q^n, \quad q^{n+1} = r^{n+1} - \frac{(Ar^{n+1}, q^n)}{(Aq^n, q^n)} q^n, \quad U^0 = 0,\ q^0 = b.$$

In the preconditioned conjugate gradient (PCG) method the conjugate gradient method is applied to equation (B.1) after multiplication by some SPD approximation B of A^{-1}, which is easier to determine than A^{-1}, so that the equation (B.1) may be written $BAU = Bb$. We note that BA is SPD with respect to the inner product $(B^{-1}V, W)$. The error estimate (B.5) is now valid in the corresponding norm with $\kappa = \kappa(BA)$; B would be chosen so that this condition number is smaller than $\kappa(A)$. For the recursion formulas the only difference is that now $r^n = B(b - AU^n)$ and $q^0 = Bb$.

B.5 Multigrid and Domain Decomposition Methods

In the case that the system (B.1) comes from a standard finite element problem, one way of defining a preconditioner as an approximate inverse of A is by means of the multigrid method. This method is based on the observation that large components of the errors are associated with low frequencies in a spectral representation. The basic idea is then to work in a systematic way with a sequence of triangulations and to reduce the low frequency errors on coarse triangulations, which corresponds to small size problems, and to reduce the higher frequency, or oscillatory, residual errors on finer triangulations by a smoothing operator, such as a step of the Jacobi method, which is relatively inexpensive.

Assuming that Ω is a plane polygonal domain we may, for instance, proceed as follows. We first perform a coarse triangulation of Ω. Each of the triangles is then divided into four similar triangles, and this process is repeated, which after a finite number M of steps leads to a fine triangulation with each of the original triangles divided into 4^M small triangles. It is on this fine triangulation which we want to use the finite element method, and thus to define an iterative method. To find the next iterate U^{n+1} from U^n we start at the finest triangulation and go recursively from one level of fineness to the previous in three steps:

1. A preliminary smoothing on the finer of the present triangulations.
2. Correction on the coarser triangulation by solving a residual equation.
3. A postsmoothing on the finer triangulation.

The execution of step 2 is thus itself carried out in three steps, starting with a smoothing on the present level and going to step 2 on the next coarser level, until one arrives at the original coarse triangulation, where the corresponding residual equation is solved exactly. Postsmoothing on successive finer levels then completes the algorithm for computing the next iterate U^{n+1}. This particular procedure is referred to as the V-cycle algorithm. It turns out that, under the appropriate assumptions, the error reduction matrix R satisfies $\|R\| \leq \rho < 1$, with ρ independent of M, i.e., of h, and that the number of operations is of order $O(N)$, where $N = O(h^{-2})$ is the dimension of the matrix associated with the finest triangulation.

A class of iterative methods that have attracted a lot of attention recently is the so called domain decomposition methods. These assume that the domain Ω in which we want to solve our elliptic problem may be decomposed into subdomains $\Omega_j,\ j = 1, \dots, M$, which could overlap. The idea is to reduce the boundary value problem on Ω into problems on each of the Ω_j, which are then coupled by their values on the intersections. The problems on the Ω_j could be solved independently on parallel processors. This is particularly efficient when the individual problems may be solved very fast, e.g., by fast transform methods.

The domain decomposition methods go back to the Schwarz alternating procedure, in which $\Omega = \Omega_1 \cup \Omega_2$ for two overlapping domains Ω_1 and Ω_2. Considering the Dirichlet problem (1.1) and (1.2) on Ω (with $g = 0$ on Γ) one defines a sequence $\{u^k\}_{k=0}^{\infty}$ starting with a given u^0 vanishing on $\partial\Omega$, by

$$
\begin{aligned}
-\Delta u^{2k+1} &= f \qquad \text{in } \Omega_1, \\
u^{2k+1} &= \begin{cases} u^{2k} & \text{on } \partial\Omega_1 \cap \Omega_2, \\ 0 & \text{on } \partial\Omega_1 \cap \partial\Omega, \end{cases} \\
-\Delta u^{2k+2} &= f \qquad \text{in } \Omega_2, \\
u^{2k+2} &= \begin{cases} u^{2k+1} & \text{on } \partial\Omega_2 \cap \Omega_1, \\ 0 & \text{on } \partial\Omega_2 \cap \partial\Omega, \end{cases}
\end{aligned}
$$

and this procedure can be combined with numerical solution by, e.g., finite elements.

The following alternative approach may be pursued when Ω_1 and Ω_2 are disjoint but with a common interface $\partial\Omega_1 \cap \partial\Omega_2$: If u_j denotes the solution in $\Omega_j,\ j = 1, 2$, then the transmission conditions $u_1 = u_2,\ \partial u_1/\partial n = \partial u_2/\partial n$ have to be satisfied on the interface. One method is then to reduce the problem to an integral type equation on the interface and use this as a basis of an iterative method.

Bibliography

Partial Differential Equations

R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 1–6*, Springer-Verlag, Berlin, 1988–1993.

L. C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, RI, 1998.

G. B. Folland, *Introduction to Partial Differential Equations*, second ed., Princeton University Press, Princeton, NJ, 1995.

A. Friedman, *Partial Differential Equations*, Holt, Rinehart and Winston, Inc., New York, 1969.

P. R. Garabedian, *Partial Differential Equations*, AMS Chelsea Publishing, Providence, RI, 1998, Reprint of the 1964 original.

F. John, *Partial Differential Equations*, fourth ed., Springer-Verlag, New York, 1991.

I. G. Petrovsky, *Lectures on Partial Differential Equations*, Dover Publications Inc., New York, 1991, Translated from the Russian by A. Shenitzer, Reprint of the 1964 English translation.

M. H. Protter and H. F. Weinberger, *Maximum Principles in Differential Equations*, Springer-Verlag, New York, 1984, Corrected reprint of the 1967 original.

J. Rauch, *Partial Differential Equations*, Springer-Verlag, New York, 1991.

M. Renardy and R. C. Rogers, *An Introduction to Partial Differential Equations*, Springer-Verlag, New York, 1993.

Functional Analysis

L. Debnath and P. Mikusiński, *Introduction to Hilbert Spaces with Applications*, second ed., Academic Press Inc., San Diego, CA, 1999.

E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons Inc., New York, 1989.

W. Rudin, *Functional Analysis*, second ed., McGraw-Hill Inc., New York, 1991.

G. F. Simmons, *Introduction to Topology and Modern Analysis*, Robert E. Krieger Publishing Co. Inc., Melbourne, Fla., 1983.

Finite Element Methods

D. Braess, *Finite Elements*, second ed., Cambridge University Press, Cambridge, 2001.

S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, second ed., Springer-Verlag, New York, 2002.

P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Introduction to adaptive methods for differential equations*, Acta Numerica, 1995, Cambridge Univ. Press, Cambridge, 1995, pp. 105–158.

G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall Inc., Englewood Cliffs, N. J., 1973.

V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, Berlin, 1997.

O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method. Vol. 1–3*, Fifth edition, Butterworth-Heinemann, Oxford, 2000.

Finite Difference Methods

G. E. Forsythe and W. R. Wasow, *Finite-Difference Methods for Partial Differential Equations*, John Wiley & Sons Inc., New York, 1960.

B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods*, John Wiley & Sons Inc., New York, 1995.

R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1967.

J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

Other Classes of Numerical Methods

K. E. Atkinson, *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge University Press, Cambridge, 1997.

J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, second ed., Dover Publications Inc., Mineola, NY, 2001.

C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, New York, 1988.

G. Chen and J. Zhou, *Boundary Element Methods*, Computational Mathematics and Applications, Academic Press, London, 1992.

J. Douglas, Jr. and T. Dupont, *Collocation Methods for Parabolic Equations in a Single Space Variable*, Springer-Verlag, Berlin, 1974, Lecture Notes in Mathematics, Vol. 385.

D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1977.

R. Li, Z. Chen, and W. Wu, *Generalized Difference Methods for Differential Equations*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 226, Marcel Dekker Inc., New York, 2000,

A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, vol. 23, Springer-Verlag, Berlin, 1994.

L. N. Trefethen, *Spectral Methods in MATLAB*, Software, Environments, and Tools, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

W. L. Wendland, *Boundary element methods for elliptic problems*, Mathematical Theory of Finite and Boundary Element Methods (A. H. Schatz, V. Thomée, and W. L. Wendland, eds.), Birkhäuser Verlag, Basel, 1990, pp. 219–276.

Numerical Linear Algebra

J. H. Bramble, *Multigrid Methods*, Longman Scientific & Technical, Harlow, 1993.

J. W. Demmel, *Applied Numerical Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

P. Deuflhard and A. Hohmann, *Numerical Analysis in Modern Scientific Computing*, second ed., Springer, New York, 2003.

G. H. Golub and C. F. Van Loan, *Matrix Computations*, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.

A. Quarteroni and A. Valli, *Domain Decomposition Methods for Partial Differential Equations*, Oxford University Press, New York, 1999.

B. F. Smith, P. E. Bjørstad, and W. D. Gropp, *Domain Decomposition*, Cambridge University Press, Cambridge, 1996.

L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

R. S. Varga, *Matrix Iterative Analysis*, expanded ed., Springer-Verlag, Berlin, 2000.

Index

Texts in Applied Mathematics

(continued from page ii)

34. *Chicone:* Ordinary Differential Equations with Applications, Second Edition.
35. *Kevorkian:* Partial Differential Equations: Analytical Solution Techniques, Second Edition.
36. *Dullerud/Paganini:* A Course in Robust Control Theory: A Convex Approach.
37. *Quarteroni/Sacco/Saleri:* Numerical Mathematics.
38. *Gallier:* Geometric Methods and Applications: For Computer Science and Engineering.
39. *Atkinson/Han:* Theoretical Numerical Analysis: A Functional Analysis Framework, Second Edition.
40. *Brauer/Castillo-Chávez:* Mathematical Models in Population Biology and Epidemiology.
41. *Davies:* Integral Transforms and Their Applications, Third Edition.
42. *Deuflhard/Bornemann:* Scientific Computing with Ordinary Differential Equations.
43. *Deuflhard/Hohmann:* Numerical Analysis in Modern Scintific Computing: An Introduction, Second Edition.
44. *Knabner/Angermann:* Numerical Methods for Elliptic and Parabolic Partial Differential Equations.
45. *Larsson/Thomée:* Partial Differential Equations with Numerical Methods.
46. *Pedregal:* Introduction to Optimization.
47. *Ockendon/Ockendon:* Waves and Compressible Flow.
48. *Hinrichsen/Pritchard:* Mathematical Systems Theory I.
49. *Bullo/Lewis:* Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems.
50. *Verhulst:* Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics.
51. *Bondeson/Rylander/Ingelström:* Computational Electromagnetics.
52. *Holmes:* Introduction to Numerical Methods in Differential Equations.
53. *Pavliotis/Stuart:* Multiscale Methods: Averaging and Homogenization.
54. *Hesthaven/Warburton:* Nodal Discontinuous Galerkin Methods.
55. *Allaire/Kaber:* Numerical Linear Algebra.