Analysis of Partial Differential Equations

Lectures and Problems

Brando Magnani

Acknowledgements

This book grew out of handwritten notes taken during a PhD-level class in partial differential equations at the Courant Institute of Mathematical Sciences, New York University. The class was taught by Professor Jalal Shatah.

Contents

Chapter 1

Characteristics and the Cauchy Problem

1.1 First-order equations and characteristic curves

Consider the first-order equation

$$a(x,t)\,\partial_t u + b(x,t)\,\partial_x u = f(x,t). \tag{1.1}$$

On a region where $a \neq 0$, division by a gives

$$\partial_t u + \widetilde{b}(x,t)\,\partial_x u = \widetilde{f}(x,t), \qquad \widetilde{b} = \frac{b}{a}, \quad \widetilde{f} = \frac{f}{a}. \tag{1.2}$$

For notational simplicity, we now rename $\widetilde{b}$ and $\widetilde{f}$ as b and f.

Let $t \mapsto x(t)$ be a curve in the (x,t)-plane. Along this curve, the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(x(t),t\big) = \partial_t u\big(x(t),t\big) + \dot{x}(t)\,\partial_x u\big(x(t),t\big). \tag{1.3}$$

If $x(t)$ solves

$$\dot{x}(t) = b\big(x(t),t\big), \tag{1.4}$$

then (1.2) becomes the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(x(t),t\big) = f\big(x(t),t\big). \tag{1.5}$$

The curves determined by (1.4) are the *characteristics*. The PDE transports the value of u along these curves.

The initial curve must cross the characteristic direction. If the initial information is prescribed along one characteristic, the construction does not spread that information to a two-dimensional neighborhood, so the method does not determine a local solution there.

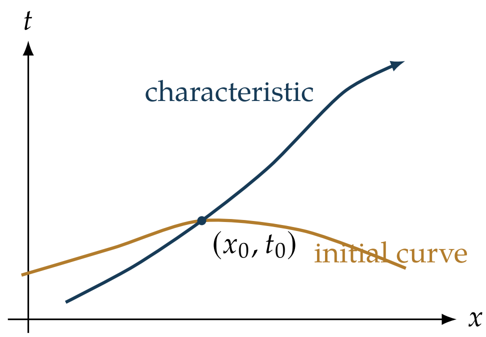


Figure 1.1: Initial data must be transverse to the characteristic direction.

1.1.1 A quasilinear equation in two variables

Consider

$$a(x,y)\,\partial_x u + b(x,y)\,\partial_y u = f(x,y,u). \tag{1.6}$$

Parametrize a curve by $s \mapsto (x(s), y(s))$. Along the curve,

$$\frac{\mathrm{d}}{\mathrm{d}s} u\big(x(s), y(s)\big) = \dot{x}(s)\,\partial_x u + \dot{y}(s)\,\partial_y u. \tag{1.7}$$

Set $z(s) = u(x(s), y(s))$. The characteristic system is

$$\begin{cases} \dot{x} = a(x,y), \\ \dot{y} = b(x,y), \\ \dot{z} = f(x,y,z). \end{cases} \tag{1.8}$$

Thus z records the value of the solution along each characteristic, and the first-order PDE becomes a system of ODEs.

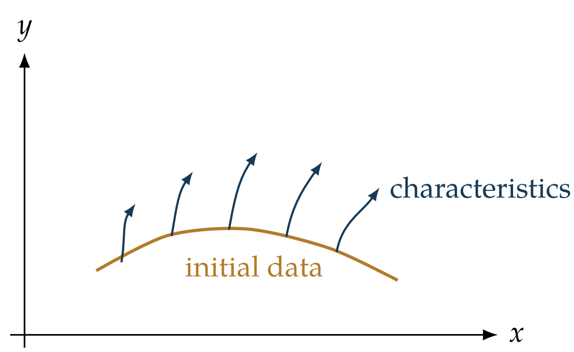


Figure 1.2: A transversal initial curve launches a family of characteristics.

1.1.2 Multi-index notation and several variables

Write

$$x = (x^1, x^2, \dots, x^m) \in \mathbb{R}^m, \qquad \partial_i u := \partial_{x^i} u.$$

We use the Einstein summation convention: a repeated index is summed. For example,

$$a^i \partial_i u = \sum_{i=1}^{m} a^i \partial_i u.$$

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$, write

$$x^\alpha = (x^1)^{\alpha_1} \cdots (x^m)^{\alpha_m}, \qquad \partial^\alpha u = \partial_1^{\alpha_1} \cdots \partial_m^{\alpha_m} u.$$

Now consider the first-order quasilinear equation

$$a^i(x, u)\, \partial_i u = f(x, u), \tag{1.9}$$

where $a = (a^1, \dots, a^m)$. Let $s \mapsto \gamma(s)$ be a characteristic and put $z(s) = u(\gamma(s))$. Since

$$\dot z(s) = \dot\gamma^i(s)\, \partial_i u\big(\gamma(s)\big),$$

the characteristic system is

$$\begin{cases} \dot\gamma^i = a^i(\gamma, z), \quad i = 1, \dots, m, \\ \dot z = f(\gamma, z), \end{cases} \qquad \gamma(0) = x_0, \quad z(0) = z_0. \tag{1.10}$$

This is an initial-value problem for an ODE in $m + 1$ unknowns.

If a and f are locally Lipschitz in the relevant variables, this ODE system has a unique local solution. With the corresponding smooth dependence on initial data, a transversal initial hypersurface supplies local coordinates for the characteristic flow and gives a local solution of (1.9).

1.2 The ODE result behind the characteristic method

The characteristic construction uses local existence and uniqueness for autonomous ODEs. We now prove the result used above and record its continuation principle.

1.2.1 Local existence by Picard iteration

Consider

$$\dot x = F(x), \qquad x(0) = x_0, \qquad x \in \mathbb{R}^d, \tag{1.11}$$

where F is locally Lipschitz. A locally Lipschitz map is Lipschitz on each compact set. In particular, on the closed ball $\overline{B}_2(x_0)$ there are finite constants

$$M = \sup_{x \in \overline{B}_2(x_0)} |F(x)|, \qquad \big|F(x) - F(y)\big| \le L\,\big|x - y\big| \quad (x, y \in \overline{B}_2(x_0)).$$

Let

$$X = C\big([0,T], \overline{B}_2(x_0)\big), \qquad \|x\|_\infty = \sup_{0\le t\le T} |x(t)|\,,$$

and define

$$(\Phi x)(t) = x_0 + \int_0^t F(x(s))\,\mathrm{d}s. \tag{1.12}$$

Choose $T > 0$ so that

$$TM \le 1, \qquad TL = q < \frac{1}{2}. \tag{1.13}$$

Then

$$|(\Phi x)(t) - x_0| \le \int_0^t |F(x(s))|\,\mathrm{d}s \le TM \le 1,$$

so Φ maps X into itself. Moreover,

$$\begin{aligned} \left\|\Phi x - \Phi y\right\|_\infty &\le \sup_{0\le t\le T} \int_0^t \left|F(x(s)) - F(y(s))\right|\,\mathrm{d}s \qquad (1.14)\\ &\le TL\left\|x - y\right\|_\infty = q\left\|x-y\right\|_\infty. \qquad (1.15) \end{aligned}$$

The set X is closed in the Banach space $C([0,T],\mathbb{R}^d)$, hence it is complete, and Φ is a contraction. For the Picard iterates $x^{(n+1)} = \Phi x^{(n)}$, the contraction estimate also gives

$$\begin{aligned} \left\|x^{(n+1)} - x_0\right\|_\infty &\le \sum_{k=0}^{n} \left\|x^{(k+1)} - x^{(k)}\right\|_\infty \\ &\le \sum_{k=0}^{n} q^k \left\|x^{(1)} - x^{(0)}\right\|_\infty \le \frac{1}{1-q}\left\|x^{(1)} - x^{(0)}\right\|_\infty. \end{aligned}$$

When $q < 1/2$ and the first step lies in $\overline{B}_1(x_0)$, the iterates remain in $\overline{B}_2(x_0)$. Banach's fixed-point theorem therefore gives a unique fixed point of Φ. This fixed point is the unique local solution of (1.11).

Theorem 1.1 (Local existence and uniqueness for an autonomous ODE)**.** *If $F : \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz, then for every $x_0 \in \mathbb{R}^d$ there is a time $T > 0$ for which* (1.11) *has a unique solution on* $[0,T]$.

1.2.2 Continuation and global existence

Suppose a solution of $\dot{x} = F(x)$ is known on $[0,T]$, and write $x_T = x(T)$. For $t \ge T$, restart the integral equation at x_T:

$$(\Phi_T x)(t) = x_T + \int_T^t F(x(s))\,\mathrm{d}s. \tag{1.16}$$

The same contraction argument extends the solution to $[T, T+T_1]$. If the trajectory stays in a fixed compact set, the bound and Lipschitz constant for F can be chosen uniformly there, so T_1 can be chosen independently of T. Repeating the construction then extends the solution for all positive times.

The general continuation statement is most naturally written for a possibly nonautonomous equation. Let $D \subset \mathbb{R} \times \mathbb{R}^m$ be open and consider

$$\dot{x} = f(t, x), \qquad x(t_0) = x_0, \tag{1.17}$$

where f is continuous and locally Lipschitz in x.

Theorem 1.2 (Continuation criterion)**.** *Let (T_-, T_+) be the maximal interval of existence of the solution of (1.17). If $T_+ < \infty$, then the graph $(t, x(t))$ leaves every compact subset of D as $t \uparrow T_+$. More precisely, for every compact $K \subset D$ there is $T \in (t_0, T_+)$ such that*

$$(t, x(t)) \notin K \qquad \text{for every } t \in (T, T_+).$$

For an autonomous system on an open set $U \subset \mathbb{R}^m$, the contrapositive is useful: if $x(t)$ stays in a compact set $C \subset U$ for every $t < T_+$, then $T_+ = \infty$.

Example 1.3 (An a priori bound prevents finite-time breakdown)**.** Suppose an autonomous solution on $\mathbb{R}^m$ satisfies

$$|x(t)| \le \mathrm{e}^t\, |x(0)| \qquad (0 \le t < T_+).$$

If $T_+ < \infty$, then the whole trajectory lies in the compact ball

$$\left\{x \in \mathbb{R}^m : |x| \le \mathrm{e}^{T_+}\, |x(0)|\right\}.$$

This contradicts the continuation criterion. Hence $T_+ = \infty$.

The same idea appears in evolution equations. The notes use the schematic damped semilinear wave equation

$$\partial_t^2 u - \Delta u + \partial_t u + u^3 = 0 \tag{1.18}$$

as an example: a finite-time bound on the phase-space norm, of the form

$$\|(u, \partial_t u)(t)\|_{H^1 \times L^2} \le C_T\, \|(u, \partial_t u)(0)\|_{H^1 \times L^2},$$

keeps the solution in a bounded region of the local theory and allows the argument to restart. On a short interval after time T, the corresponding control quantity y is estimated by $y(t) \le 2y(T)$; repeating this short-time estimate is the PDE analogue of the ODE continuation argument.

1.3 Second-order equations and Cauchy data

Continuity of the characteristic vector field is enough for an ODE existence theorem, while local Lipschitz regularity is used for uniqueness. A classical first-order PDE asks only for a C^1 solution. For a second-order PDE, however, one must also produce second derivatives.

Consider

$$a(x, y)\, \partial_x^2 u + 2b(x, y)\, \partial_{xy} u + c(x, y)\, \partial_y^2 u = f(x, y). \tag{1.19}$$

In index notation, the principal part has the form

$$a^{ij}(x)\, \partial_{ij} u = f(x). \tag{1.20}$$

If u is prescribed on a hypersurface, its trace determines all tangential derivatives there, but it does not determine transverse derivatives. A second-order Cauchy problem therefore prescribes the value of u and one transverse first derivative.

Let the initial curve be

$$\Sigma = \left\{(x, y) : \phi(x, y) = 0\right\}.$$

Near a point where the coordinate change is nonsingular, introduce

$$t = \phi(x, y), \qquad s = x, \qquad u(x, y) = \overline{u}\big(\phi(x, y), x\big).$$

The chain rule gives

$$\partial_x u = (\partial_x \phi)\, \partial_t \overline{u} + \partial_s \overline{u}. \tag{1.21}$$

The datum $\overline{u}(0, s)$ determines all tangential derivatives. Once $\partial_t \overline{u}(0, s)$ is also prescribed, the complete first-order derivative data of u on Σ are known.

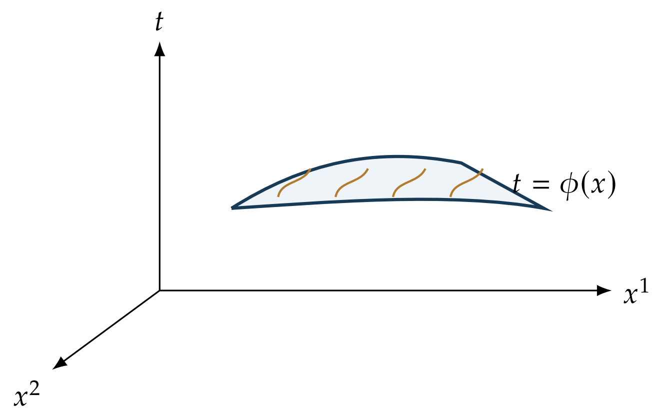


Figure 1.3: Cauchy data on a hypersurface $t = \phi(x)$. Tangential derivatives are determined by the trace; a transverse derivative must be prescribed separately.

1.3.1 The model initial curve $y = 0$

Take $\phi(x, y) = y$, so that $\Sigma = \left\{y = 0\right\}$. Prescribe

$$u(x, 0) = g_0(x), \qquad \partial_y u(x, 0) = g_1(x). \tag{1.22}$$

Tangential differentiation gives

$$\partial_x^2 u(x, 0) = g_0''(x), \qquad \partial_{xy} u(x, 0) = g_1'(x).$$

The only second derivative not yet determined is $\partial_y^2 u(x, 0)$. Restricting (1.19) to $y = 0$ gives

$$a(x, 0)\, \partial_x^2 u(x, 0) + 2b(x, 0)\, \partial_{xy} u(x, 0) + c(x, 0)\, \partial_y^2 u(x, 0) = f(x, 0). \tag{1.23}$$

If $c(x, 0) \neq 0$, then

$$\partial_y^2 u(x, 0) = \frac{f(x, 0) - a(x, 0) g_0''(x) - 2b(x, 0) g_1'(x)}{c(x, 0)}. \tag{1.24}$$

Repeated differentiation of the PDE determines all higher transverse derivatives, provided the coefficients and data have enough regularity. Thus $y = 0$ is noncharacteristic where $c(x, 0) \neq 0$ and characteristic where $c(x, 0) = 0$.

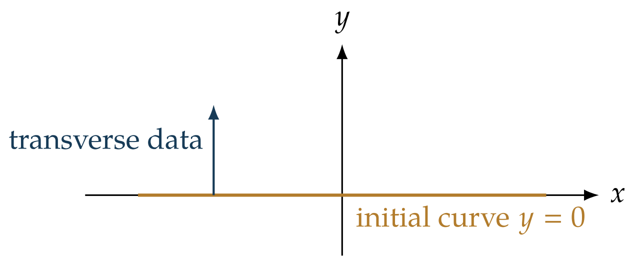


Figure 1.4: On $y = 0$, the trace gives tangential derivatives and the Cauchy problem supplies one transverse first derivative.

1.4 Characteristic hypersurfaces in several variables

Let $x \in \mathbb{R}^m$ and consider

$$a^{ij}(x)\,\partial_{ij}u = f(x) \tag{1.25}$$

with Cauchy data on

$$\Sigma = \left\{x : \phi(x) = 0\right\}, \qquad \nabla\phi \neq 0 \quad \text{on } \Sigma.$$

Choose adapted local coordinates

$$x \longleftrightarrow (t, s^1, \ldots, s^{m-1}), \qquad t = \phi(x),$$

and write $u(x) = \overline{u}(t, s)$. The Cauchy data may be written as

$$\overline{u}|_{t=0} = u_0, \qquad \partial_t\overline{u}|_{t=0} = u_1. \tag{1.26}$$

Equivalently, one may prescribe a derivative in any nonzero transverse direction, such as $\nabla\phi \cdot \nabla u$; the trace and its tangential derivatives then determine the conversion to $\partial_t\overline{u}$.

Write

$$\phi_i = \partial_i\phi, \qquad s_i^k = \partial_i s^k.$$

The first derivatives transform as

$$\partial_i u = \phi_i\,\partial_t\overline{u} + s_i^k\,\partial_{s^k}\overline{u}. \tag{1.27}$$

Differentiating once more gives

$$\begin{aligned}\partial_{ij}u = {} & \phi_i\phi_j\,\partial_t^2\overline{u} + (\partial_{ij}\phi)\,\partial_t\overline{u} + \phi_i s_j^k\,\partial_{ts^k}\overline{u} + \phi_j s_i^k\,\partial_{ts^k}\overline{u} \\ & + (\partial_{ij}s^k)\,\partial_{s^k}\overline{u} + s_i^k s_j^\ell\,\partial_{s^k s^\ell}\overline{u}.\end{aligned} \tag{1.28}$$

On $t = 0$, every term on the right except $\partial_t^2\overline{u}$ is already known from the Cauchy data and their tangential derivatives. Substitution into (1.25) gives

$$\underbrace{a^{ij}(x)\,\partial_i\phi(x)\,\partial_j\phi(x)}_{Q(x,\nabla\phi)}\,\partial_t^2\overline{u} + \text{known terms} = f(x). \tag{1.29}$$

Definition 1.4. The hypersurface $\Sigma = \{\phi = 0\}$ is *noncharacteristic* at x when

$$Q(x, \nabla\phi) = a^{ij}(x)\,\partial_i\phi(x)\,\partial_j\phi(x) \neq 0. \tag{1.30}$$

It is *characteristic* where this quantity is zero.

For the model curve $y = 0$, one has $\phi(x, y) = y$ and $Q(x, \nabla\phi) = c(x, y)$, which recovers the criterion above. When $Q \neq 0$, the PDE determines $\partial_t^2 \overline{u}$ on the initial hypersurface. Repeated differentiation determines all higher transverse derivatives recursively.

At order $|\alpha|$, the product rule takes the form

$$\partial^\alpha(FG) = F\,\partial^\alpha G + G\,\partial^\alpha F + \sum_{\substack{\beta\le\alpha \\ \beta\neq 0,\ \beta\neq\alpha}} \binom{\alpha}{\beta}(\partial^\beta F)(\partial^{\alpha-\beta}G). \tag{1.31}$$

Thus all terms other than the highest transverse derivative involve quantities that have already been found. This produces a formal Taylor series in the transverse variable.

1.5 The Cauchy–Kowalevski theorem

The preceding calculation produces formal Taylor coefficients. The analytic hypotheses in the Cauchy–Kowalevski theorem guarantee that the resulting series converges; smoothness alone does not provide this conclusion.

Theorem 1.5 (Cauchy–Kowalevski, second-order form)**.** *Let the coefficients a^{ij}, the right-hand side f, the defining function ϕ, and the Cauchy data u_0, u_1 be analytic near a point $x_0 \in \Sigma$. If $\Sigma = \{\phi = 0\}$ is noncharacteristic at x_0, then*

$$a^{ij}(x)\,\partial_{ij}u = f(x), \qquad u|_\Sigma = u_0, \qquad \partial_t\overline{u}|_\Sigma = u_1$$

has a unique analytic solution in a neighborhood of x_0.

More generally, an analytic PDE with analytic Cauchy data on an analytic noncharacteristic hypersurface has a unique local analytic solution once it is solved for its highest transverse derivative. Uniqueness comes from the fact that the PDE and data determine every transverse Taylor coefficient.

1.5.1 First-order systems

Let

$$\mathbf{u} = (u_1, \ldots, u_N)^\mathsf{T}$$

and consider

$$A^i(x)\,\partial_i\mathbf{u} + B(x)\mathbf{u} = F(x), \tag{1.32}$$

where the A^i and B are matrices. For $\Sigma = \{\phi = 0\}$, define

$$A(\nabla\phi) = A^i\partial_i\phi. \tag{1.33}$$

The hypersurface is noncharacteristic precisely when

$$\det\big(A(\nabla\phi)\big) \neq 0. \tag{1.34}$$

Under analytic hypotheses, Cauchy–Kowalevski gives a unique local analytic solution.

1.5.2 Higher-order scalar equations

For a scalar equation of order k with principal part

$$a^{i_1\cdots i_k}(x)\partial_{i_1\cdots i_k}u + \text{lower-order terms} = 0, \tag{1.35}$$

the characteristic condition for $\Sigma = \big\{\phi = 0\big\}$ is

$$a^{i_1\cdots i_k}(x)\partial_{i_1}\phi(x)\cdots\partial_{i_k}\phi(x) = 0. \tag{1.36}$$

Only the highest-order part of the equation and the normal covector $\nabla\phi$ enter this condition.

Chapter 2

Cauchy Problems, Principal Symbols, and Harmonic Functions

The theory of ordinary differential equations suggests how to formulate initial-value problems for partial differential equations. For a PDE, the initial data are prescribed on a hypersurface rather than at a point. Whether the equation can propagate those data depends on its highest-order terms and on the geometry of the hypersurface. This leads to characteristic surfaces, the principal symbol, and the elliptic–hyperbolic–parabolic classification. The second half of the chapter begins the study of the elliptic model operator, the Laplacian, and derives the mean-value property of harmonic functions.

2.1 Ordinary differential equations as a model

Consider the autonomous initial-value problem

$$\dot{x} = F(x), \qquad x(0) = x_0, \qquad x \in \mathbb{R}^m. \tag{2.1}$$

The basic local theory is as follows.

1. If F is continuous, then a local solution exists.
2. If F is locally Lipschitz in x, then that local solution is unique.

An ordinary differential equation of order ℓ,

$$\sum_{k=0}^{\ell} a_k x^{(k)}(t) = f(t), \qquad a_\ell \neq 0, \tag{2.2}$$

can be written as a first-order system. Set

$$y_1 = x, \qquad y_2 = x^{(1)}, \qquad \dots, \qquad y_\ell = x^{(\ell-1)}. \tag{2.3}$$

Then (2.2) becomes a system

$$\dot{y} = G(t, y). \tag{2.4}$$

Thus an equation of order ℓ requires ℓ pieces of initial data,

$$x(0),\ x^{(1)}(0),\ \dots,\ x^{(\ell-1)}(0). \tag{2.5}$$

2.2 The Cauchy problem for a PDE

For a PDE of order k, a *Cauchy problem* prescribes the unknown and its first $k-1$ transverse derivatives on an initial hypersurface and asks for a local solution. In coordinates in which the initial hypersurface is $t=0$, the k pieces of data take the form

$$\partial_t^j u\Big|_{t=0} = u_j, \qquad j = 0, \dots, k-1. \tag{2.6}$$

Typical second-order examples are

$$\partial_x^2 u = f(x), \qquad \text{an ordinary differential equation,} \tag{2.7}$$
$$\partial_x^2 u + \partial_y^2 u = f(x), \tag{2.8}$$
$$\partial_{xy} u = f(x,y). \tag{2.9}$$

For either PDE, the schematic second-order data on a smooth initial curve Γ are

$$u|_\Gamma = u_0, \qquad \partial_\nu u|_\Gamma = u_1, \tag{2.10}$$

where ν is a chosen direction transverse to Γ. The direction need not be a unit normal; what matters here is transversality.

The power-series idea is already visible for the ODE

$$x'' + ax' + bx = f(t), \qquad x(0) = x_0, \qquad x'(0) = x_1. \tag{2.11}$$

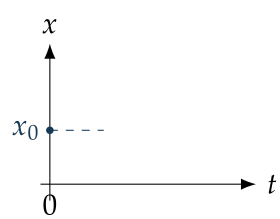


Figure 2.1: The Cauchy data are prescribed at the initial time $t=0$.

The equation determines

$$x''(0) = f(0) - ax_1 - bx_0. \tag{2.12}$$

Differentiating the equation determines $x'''(0)$, and repetition determines all higher derivatives at the origin. If the coefficients and forcing are analytic, these derivatives produce the Taylor series

$$x(t) = \sum_{r=0}^{\infty} \frac{x^{(r)}(0)}{r!} t^r. \tag{2.13}$$

The same recursive idea motivates the analytic Cauchy problem for PDEs.

2.3 Characteristic and noncharacteristic data

Return to the model equation

$$\partial_{xy} u = f(x, y) \tag{2.14}$$

with Cauchy data prescribed on a smooth curve Γ. Knowing u on the curve determines all derivatives tangent to the curve. To move off the curve, one also prescribes one transverse derivative.

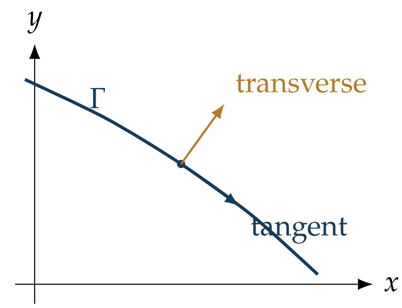


Figure 2.2: Data on Γ determine tangential derivatives; one transverse derivative must be prescribed separately.

Write the curve locally as

$$\Gamma = \{(x, y) : \phi(x, y) = 0\}, \qquad d\phi \neq 0 \quad \text{on } \Gamma. \tag{2.15}$$

Choose a second function g so that

$$t = \phi(x, y), \qquad z = g(x, y) \tag{2.16}$$

is a local coordinate system. Thus t is transverse and z is tangential. Write

$$u(x, y) = \widetilde{u}(t, z). \tag{2.17}$$

The Cauchy data become

$$\widetilde{u}(0, z) = u_0(z), \qquad \partial_t \widetilde{u}(0, z) = u_1(z). \tag{2.18}$$

Consequently, the data determine

$$\partial_z^r \widetilde{u}|_{t=0} \quad \text{and} \quad \partial_z^r \partial_t \widetilde{u}|_{t=0}\,, \qquad r = 0, 1, 2, \ldots. \tag{2.19}$$

The first unknown pure transverse derivative is $\partial_t^2 \widetilde{u}\big|_{t=0}$.

The chain rule gives

$$\partial_x u = (\partial_x \phi)\partial_t \widetilde{u} + (\partial_x z)\partial_z \widetilde{u}, \tag{2.20}$$
$$\partial_y u = (\partial_y \phi)\partial_t \widetilde{u} + (\partial_y z)\partial_z \widetilde{u}. \tag{2.21}$$

Differentiating once more,

$$\partial_{xy} u = (\partial_x \phi)(\partial_y \phi)\partial_t^2 \widetilde{u} + K, \tag{2.22}$$

where

$$K = (\partial_{xy}\phi)\partial_t\widetilde{u} + (\partial_x\phi)(\partial_y z)\partial_{tz}\widetilde{u} + (\partial_x z)(\partial_y\phi)\partial_{zt}\widetilde{u} + (\partial_{xy}z)\partial_z\widetilde{u} + (\partial_x z)(\partial_y z)\partial_z^2\widetilde{u}. \tag{2.23}$$

Every term in K is already known on $t = 0$. Hence (2.14) determines $\partial_t^2\widetilde{u}$ when

$$\boxed{(\partial_x\phi)(\partial_y\phi) \neq 0.} \tag{2.24}$$

When the product vanishes, Γ is characteristic for this equation; when it does not vanish, Γ is noncharacteristic. In the analytic setting, one may differentiate the PDE again to determine the third transverse derivative and continue recursively to construct the local power series.

2.4 Linear operators and the principal symbol

Let $x = (x_1, \ldots, x_m)$, let $I = (i_1, \ldots, i_m)$ be a multi-index, and write

$$|I| = i_1 + \cdots + i_m, \qquad \partial^I = \partial_1^{i_1} \cdots \partial_m^{i_m}. \tag{2.25}$$

A linear scalar differential operator of order k has the form

$$L(x,\partial)u = \sum_{|I|\leq k} a_I(x)\partial^I u, \qquad L(x,\partial)u = f(x). \tag{2.26}$$

Its highest-order, or *principal*, part is

$$L_k(x,\partial)u = \sum_{|I|=k} a_I(x)\partial^I u. \tag{2.27}$$

For $k = 2$, the operator may be written

$$L(x,\partial)u = \sum_{i,j=1}^{m} a^{ij}(x)\partial_{ij}u + \sum_{r=1}^{m} b^r(x)\partial_r u + c(x)u, \tag{2.28}$$

where the principal coefficients may be taken symmetric: $a^{ij} = a^{ji}$.

Introduce coordinates adapted to a hypersurface,

$$t = \phi(x), \qquad y^r = g^r(x), \qquad r = 1, \ldots, m-1. \tag{2.29}$$

Then

$$\partial_i u = (\partial_i\phi)\partial_t\widetilde{u} + \sum_{r=1}^{m-1}(\partial_i g^r)\partial_{y^r}\widetilde{u}. \tag{2.30}$$

After one more differentiation, the coefficient of the new pure transverse derivative $\partial_t^2\widetilde{u}$ is

$$\sum_{i,j=1}^{m} a^{ij}(x)(\partial_i\phi)(\partial_j\phi). \tag{2.31}$$

Therefore the hypersurface $\phi = 0$ is characteristic exactly when

$$\boxed{\sum_{i,j=1}^{m} a^{ij}(x)(\partial_i \phi)(\partial_j \phi) = 0.} \tag{2.32}$$

Only the highest-order terms enter this condition.

For an operator of general order k, its principal symbol is

$$p_k(x, \xi) = \sum_{|I|=k} a_I(x) \xi^I. \tag{2.33}$$

The coefficient of $\partial_t^k \widetilde{u}$ in adapted coordinates is

$$p_k(x, d\phi) = \sum_{|I|=k} a_I(x)(\partial \phi)^I. \tag{2.34}$$

Equivalently, with coefficients indexed as $a_{i_1 \ldots i_k}$,

$$p_k(x, d\phi) = \sum_{i_1, \ldots, i_k=1}^{m} a_{i_1 \ldots i_k}(x)(\partial_{i_1} \phi) \cdots (\partial_{i_k} \phi). \tag{2.35}$$

Thus

$$\phi = 0 \text{ is characteristic} \quad \Longleftrightarrow \quad p_k(x, d\phi) = 0. \tag{2.36}$$

2.4.1 Second-order equations in two variables

Consider

$$\begin{aligned} a(x, y)\partial_x^2 u + 2b(x, y)\partial_{xy} u + c(x, y)\partial_y^2 u \\ + d(x, y)\partial_x u + e(x, y)\partial_y u + q(x, y)u = g(x, y). \end{aligned} \tag{2.37}$$

Its characteristic equation is

$$\boxed{a(\partial_x \phi)^2 + 2b(\partial_x \phi)(\partial_y \phi) + c(\partial_y \phi)^2 = 0.} \tag{2.38}$$

Assume $a \neq 0$. Solving for $\partial_x \phi$ gives

$$\partial_x \phi = \frac{-b \pm \sqrt{b^2 - ac}}{a} \partial_y \phi. \tag{2.39}$$

To test a proposed initial curve $\psi(x, y) = 0$, substitute $\partial_x \psi$ and $\partial_y \psi$ into the left side of (2.38). The curve is noncharacteristic where the result is nonzero.

The sign of $b^2 - ac$ gives the pointwise classification:

1. If $b^2 - ac < 0$, the equation is *elliptic*; it has no real characteristic direction.
2. If $b^2 - ac > 0$, the equation is *hyperbolic*; it has two distinct real characteristic directions.
3. If $b^2 - ac = 0$, the equation is *parabolic*; it has one repeated real characteristic direction.

2.4.2 Change of variables and canonical form

The matrix formulation makes the higher-dimensional classification precise. Consider the pure second-order equation

$$\sum_{i,j=1}^{m} a_{ij}\partial_{x_i x_j} u = f, \qquad A = (a_{ij}) = A^{\mathsf{T}} \in \mathbb{R}^{m\times m}. \tag{2.40}$$

Let $B = (b_{ij})$ be invertible, set $y = Bx$, and define

$$\overline{u}(y) = u(B^{-1}y), \qquad \text{so that} \qquad u(x) = \overline{u}(Bx). \tag{2.41}$$

The chain rule gives

$$\partial_{x_k} u = \sum_i b_{ik}\partial_{y_i}\overline{u}, \tag{2.42}$$

$$\partial_{x_k x_\ell} u = \sum_{i,j} b_{ik} b_{j\ell}\partial_{y_i y_j}\overline{u}. \tag{2.43}$$

Consequently, the transformed principal matrix is

$$\overline{A} = BAB^{\mathsf{T}}, \tag{2.44}$$

and its quadratic principal symbol is

$$\xi^{\mathsf{T}}\overline{A}\xi = \xi^{\mathsf{T}}BAB^{\mathsf{T}}\xi. \tag{2.45}$$

Because A is real and symmetric, choose the rows of B to be an orthonormal eigenbasis of A. Then B is orthogonal and

$$BAB^{\mathsf{T}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_m), \tag{2.46}$$

so (2.40) takes the canonical form

$$\sum_{i=1}^{m} \lambda_i \partial_{y_i}^2 \overline{u} = f(B^{-1}y). \tag{2.47}$$

There are no mixed second derivatives in these coordinates.

The eigenvalue signs classify the principal part:

1. It is *elliptic* when every eigenvalue is nonzero and all have the same sign.
2. It is *hyperbolic* in the basic Lorentzian case when every eigenvalue is nonzero and exactly one has the sign opposite to all the others.
3. The basic *parabolic* model has one zero eigenvalue and all remaining eigenvalues have the same sign.

If there are at least two positive and at least two negative eigenvalues, the operator is often called *ultrahyperbolic*. More than one zero eigenvalue, or a zero together with mixed nonzero signs, gives a more degenerate principal part and requires a finer description.

For example,

$$\partial_t u = Lu + f, \tag{2.48}$$

with L elliptic in the spatial variable x, has no second derivative in t. The full second-order principal matrix therefore has one zero eigenvalue, corresponding to t, and eigenvalues of one sign in the spatial directions.

For constant A, the orthogonal change of variables produces the canonical form globally. A variable matrix $A(x)$ can be diagonalized algebraically at each point, but the resulting eigenbases need not form a global coordinate system; an x-dependent change of coordinates also creates lower-order terms.

Remark 2.1 (Analyticity). The recursive construction above is a power-series construction. The Cauchy–Kowalevski theorem requires analytic coefficients and forcing, an analytic initial hypersurface, analytic Cauchy data, and a noncharacteristic hypersurface. There is no unrestricted analogue with merely C^∞ data and coefficients.

2.5 The Laplacian and harmonic functions

For $u \in C^2(\Omega)$, where $\Omega \subset \mathbb{R}^m$ is open, the Laplacian is

$$\Delta u = \sum_{i=1}^{m} \partial_i^2 u. \tag{2.49}$$

The equation

$$\Delta u = f \tag{2.50}$$

is Poisson's equation. Its homogeneous case $\Delta u = 0$ is Laplace's equation.

Definition 2.2 (Harmonic function). A function $u \in C^2(\Omega)$ is *harmonic* in Ω if

$$\Delta u = 0 \qquad \text{in } \Omega. \tag{2.51}$$

2.5.1 Integration by parts and duality

Let Ω be a bounded domain with smooth boundary, let ν be the outward unit normal, and let $a = (a^1, \dots, a^m)$ be a vector field. The multidimensional integration-by-parts formula is

$$\int_\Omega a^i \partial_i f \, dx = -\int_\Omega (\partial_i a^i) f \, dx + \int_{\partial\Omega} \nu_i a^i f \, dS, \tag{2.52}$$

where repeated indices are summed. Equivalently,

$$\int_\Omega a \cdot \nabla f \, dx = -\int_\Omega (\operatorname{div} a) f \, dx + \int_{\partial\Omega} (\nu \cdot a) f \, dS. \tag{2.53}$$

Duality can be used to recognize an L^2 function. If $f \in L^1_{\text{loc}}(\Omega)$ and there is a constant C such that

$$\left| \int_\Omega f g \, dx \right| \le C \left\| g \right\|_{L^2(\Omega)} \qquad \text{for every } g \in C_c^\infty(\Omega), \tag{2.54}$$

then $f \in L^2(\Omega)$. Indeed, the left side defines a bounded linear functional on a dense subspace of L^2, so the Riesz representation theorem applies. More generally, $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leq p < \infty$.

2.5.2 Green's identities

For $u, v \in C^2(\overline{\Omega})$, integration by parts gives Green's first identity:

$$\int_\Omega (\Delta u) v \, dx = -\int_\Omega \nabla u \cdot \nabla v \, dx + \int_{\partial\Omega} (\partial_\nu u) v \, dS. \tag{2.55}$$

Integrating the interior term once more gives

$$-\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega u \Delta v \, dx - \int_{\partial\Omega} u \partial_\nu v \, dS. \tag{2.56}$$

Therefore Green's second identity is

$$\int_\Omega \big((\Delta u) v - u \Delta v\big) \, dx = \int_{\partial\Omega} \big((\partial_\nu u) v - (\partial_\nu v) u\big) \, dS. \tag{2.57}$$

If u and v are harmonic in a neighborhood of the closure of a smooth subdomain A, then

$$\int_{\partial A} (\partial_\nu u) v \, dS = \int_{\partial A} (\partial_\nu v) u \, dS. \tag{2.58}$$

Proposition 2.3 (Flux identity). *Let $A \Subset \Omega$ be a bounded domain with smooth boundary. If u is harmonic in Ω, then*

$$\int_{\partial A} \partial_\nu u \, dS = 0. \tag{2.59}$$

Proof. Set $v \equiv 1$ in Green's first identity, or apply the divergence theorem to ∇u. □

2.6 Polar coordinates and radial harmonic functions

In two-dimensional polar coordinates (r, θ),

$$\Delta u = \partial_r^2 u + \frac{1}{r} \partial_r u + \frac{1}{r^2} \partial_\theta^2 u. \tag{2.60}$$

In three-dimensional spherical coordinates (r, θ, φ), with φ the polar angle,

$$\Delta u = \partial_r^2 u + \frac{2}{r} \partial_r u + \frac{1}{r^2} \left(\partial_\varphi^2 u + cot\varphi \, \partial_\varphi u + \frac{1}{\sin^2 \varphi} \partial_\theta^2 u \right). \tag{2.61}$$

More generally, for $x \in \mathbb{R}^m \setminus \{0\}$, set

$$r = |x|, \qquad \omega = \frac{x}{|x|} \in \mathbb{S}^{m-1}. \tag{2.62}$$

Then

$$\Delta u = \partial_r^2 u + \frac{m-1}{r}\partial_r u + \frac{1}{r^2}\Delta_{\mathbb{S}^{m-1}} u, \tag{2.63}$$

where the last operator differentiates only in angular variables.

For a radial function in two dimensions, Laplace's equation becomes

$$\partial_r^2 u + \frac{1}{r}\partial_r u = 0, \tag{2.64}$$

so

$$u(r) = A \log r + B. \tag{2.65}$$

The nonconstant solution $\log r$ is harmonic for $r > 0$, but it is singular at the origin.

For a radial function in three dimensions,

$$\partial_r^2 u + \frac{2}{r}\partial_r u = 0, \tag{2.66}$$

and hence

$$u(r) = A + \frac{B}{r}. \tag{2.67}$$

In particular,

$$\Delta\left(\frac{1}{|x|}\right) = 0 \qquad \text{for } x \neq 0, \tag{2.68}$$

although $1/|x|$ is singular at the origin.

2.7 The mean-value property

2.7.1 The three-dimensional calculation

Let u be harmonic on a neighborhood of $\overline{B_R(0)} \subset \mathbb{R}^3$, and put

$$A_\varepsilon = B_R(0) \setminus \overline{B_\varepsilon(0)}, \qquad v(x) = \frac{1}{|x|}. \tag{2.69}$$

Both u and v are harmonic in A_ε.

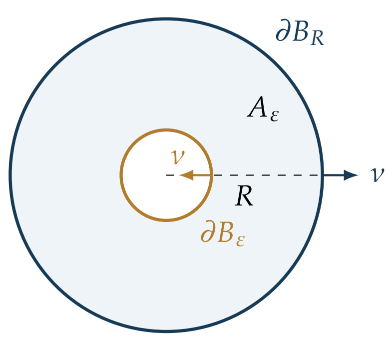


Figure 2.3: The punctured ball. On the inner boundary, the outward normal of A_ε points toward the origin.

Since

$$\partial_i v(x) = -\frac{x_i}{|x|^3}, \tag{2.70}$$

the normal derivative relative to the annulus is

$$\partial_\nu v = -\frac{1}{R^2} \quad \text{on } \partial B_R(0), \qquad \partial_\nu v = \frac{1}{\varepsilon^2} \quad \text{on } \partial B_\varepsilon(0). \tag{2.71}$$

The zero-flux identity for u on each sphere gives $\int_{\partial A_\varepsilon} v \partial_\nu u \, dS = 0$. Green's second identity therefore yields

$$\begin{aligned} 0 &= \int_{\partial A_\varepsilon} u \partial_\nu v \, dS \\ &= -\frac{1}{R^2} \int_{|x|=R} u(x) \, dS_x + \frac{1}{\varepsilon^2} \int_{|x|=\varepsilon} u(x) \, dS_x. \end{aligned} \tag{2.72}$$

Parametrizing the two spheres by $x = r\omega$, this becomes

$$0 = -\int_{\mathbb{S}^2} u(R\omega) \, dS_\omega + \int_{\mathbb{S}^2} u(\varepsilon\omega) \, dS_\omega. \tag{2.73}$$

Continuity and compactness permit $\varepsilon \downarrow 0$, giving

$$\int_{\mathbb{S}^2} u(R\omega) \, dS_\omega = 4\pi u(0). \tag{2.74}$$

2.7.2 Spherical and ball means

Let

$$\omega_m = |\mathbb{S}^{m-1}| \tag{2.75}$$

be the surface area of the unit sphere in $\mathbb{R}^m$.

Theorem 2.4 (Mean-value theorem for harmonic functions)**.** *Let $u \in C^2(\Omega)$ be harmonic in the open set $\Omega \subset \mathbb{R}^m$. If $x \in \Omega$ and $\overline{B_R(x)} \subset \Omega$, then*

$$\begin{aligned} u(x) &= \frac{1}{\omega_m} \int_{\mathbb{S}^{m-1}} u(x + R\omega) \, dS_\omega \\ &= \frac{1}{\omega_m R^{m-1}} \int_{|y|=R} u(x+y) \, dS_y. \end{aligned} \tag{2.76}$$

Integrating the spherical identity against $r^{m-1} \, dr$ gives

$$\begin{aligned} \int_{B_R(z)} u(y) \, dy &= \int_0^R \int_{|y-z|=r} u(y) \, dS_y \, dr \\ &= \int_0^R \omega_m r^{m-1} u(z) \, dr \\ &= |B_R(z)| u(z). \end{aligned} \tag{2.77}$$

Therefore

$$u(z) = \frac{1}{|B_R(z)|} \int_{B_R(z)} u(y) \, dy. \tag{2.78}$$

2.7.3 The mean-value property implies smoothness

Fix $x_0 \in \Omega$, choose $\rho > 0$ with $\overline{B_{2\rho}(x_0)} \subset \Omega$, and choose a radial bump function

$$\varphi(z) = \psi(|z|) \in C_c^\infty(B_\rho(0)) \tag{2.79}$$

normalized by

$$1 = \int_{\mathbb{R}^m} \varphi(z)\, dz = \omega_m \int_0^\rho \psi(r) r^{m-1}\, dr. \tag{2.80}$$

For $x \in B_\rho(x_0)$, polar integration and the spherical mean-value identity give

$$\begin{aligned}
(u * \varphi)(x) &= \int u(y)\varphi(x-y)\, dy \\
&= \int_{B_\rho(0)} u(x-z)\varphi(z)\, dz \\
&= \int_0^\rho \psi(r) r^{m-1} \left(\int_{\mathbb{S}^{m-1}} u(x - r\omega)\, dS_\omega \right) dr \\
&= \omega_m u(x) \int_0^\rho \psi(r) r^{m-1}\, dr = u(x).
\end{aligned} \tag{2.81}$$

The convolution $u * \varphi$ is smooth. Since this argument is local at every point, every harmonic function is in fact in $C^\infty(\Omega)$.

2.7.4 Polar volume elements and geometric constants

In two dimensions,

$$dx = r\, dr\, d\theta. \tag{2.82}$$

In m dimensions,

$$dx = r^{m-1}\, dr\, dS_\omega. \tag{2.83}$$

Consequently,

$$|\partial B_r(0)| = \omega_m r^{m-1}, \qquad \omega_2 = 2\pi, \qquad \omega_3 = 4\pi. \tag{2.84}$$

If

$$\alpha_m = |B_1(0)|, \tag{2.85}$$

then

$$\alpha_m = \frac{\omega_m}{m}, \qquad \alpha_2 = \pi, \qquad \alpha_3 = \frac{4\pi}{3}. \tag{2.86}$$

Chapter 3

Mean-Value Methods, Fundamental Solutions, and Singular Integrals

This chapter develops a sequence of consequences of the mean-value property: Liouville's theorem, the maximum principle, and uniqueness for the Dirichlet problem. It then introduces the fundamental solution of the Laplacian, principal-value distributions, singular integral operators, and the first Schauder regularity estimate. These ideas prepare the study of Dirichlet and Neumann problems and, more generally, potential theory.

3.1 The mean-value property and its consequences

Let $\Omega \subset \mathbb{R}^m$ be open. A continuous function $u\colon \Omega \to \mathbb{R}$ has the *mean-value property* if

$$u(x) = \frac{1}{|B(x,r)|} \int_{B(x,r)} u(y)\, dy \tag{3.1}$$

whenever $\overline{B(x,r)} \subset \Omega$.

Theorem 3.1 (Converse to the mean-value theorem)**.** *If $u \in C^0(\Omega)$ has the mean-value property, then $u \in C^\infty(\Omega)$ and $\Delta u = 0$ in Ω.*

Proof. Let $\varphi_\varepsilon \in C_c^\infty(\mathbb{R}^m)$ be a radial mollifier supported in $B(0,\varepsilon)$, with $\int_{\mathbb{R}^m} \varphi_\varepsilon = 1$. If $\operatorname{dist}(x, \partial\Omega) > \varepsilon$, then

$$\begin{aligned}
(u * \varphi_\varepsilon)(x) &= \int_{\mathbb{R}^m} u(y)\varphi_\varepsilon(x-y)\, dy \\
&= \int_{\mathbb{R}^m} u(x-y)\varphi_\varepsilon(y)\, dy.
\end{aligned}$$

In polar coordinates, the last integral is a weighted average of the spherical means of u about x. The mean-value property makes each such mean equal to $u(x)$; hence

$$(u * \varphi_\varepsilon)(x) = u(x) \int_{\mathbb{R}^m} \varphi_\varepsilon(y)\, dy = u(x).$$

Thus u agrees locally with a smooth mollification and is therefore smooth. For a smooth function, the small-radius expansion of its spherical mean is

$$\frac{1}{|\mathbb{S}^{m-1}|}\int_{\mathbb{S}^{m-1}} u(x+r\omega)\,d\omega = u(x) + \frac{r^2}{2m}\Delta u(x) + o(r^2).$$

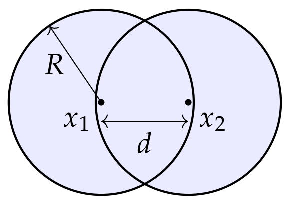
Since the left-hand side equals $u(x)$, division by r^2 followed by $r \downarrow 0$ gives $\Delta u(x) = 0$. $\square$

Theorem 3.2 (Liouville)**.** *Every bounded harmonic function on $\mathbb{R}^m$ is constant.*

Proof. Fix $x_1, x_2 \in \mathbb{R}^m$, put $d = |x_1 - x_2|$, and choose $R > d$. By the mean-value theorem,

$$\begin{aligned} |u(x_1) - u(x_2)| &\le \frac{\|u\|_{L^\infty}}{|B(0,R)|}\big|B(x_1,R) \triangle B(x_2,R)\big| \\ &\le \frac{C\|u\|_{L^\infty}}{R^m}\big((R+d)^m - (R-d)^m\big). \end{aligned}$$

The two balls and their symmetric difference are shown schematically below.



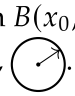
The last bound tends to zero as $R \to \infty$. Therefore $u(x_1) = u(x_2)$, and the two points were arbitrary. $\square$

We recall that $u \in C^2(\Omega)$ is harmonic precisely when $\Delta u = 0$ in Ω.

Theorem 3.3 (Strong maximum principle)**.** *Let Ω be connected and let u be harmonic in Ω. If*

$$u(x_0) = \sup_{x\in\Omega} u(x) = A$$

for some $x_0 \in \Omega$, then $u \equiv A$ in Ω.

Proof. Choose $r > 0$ with $\overline{B(x_0, r)} \subset \Omega$. The mean-value argument takes place on a ball centered at the maximizing point, . The mean-value theorem gives

$$A = u(x_0) = \frac{1}{|B(x_0,r)|}\int_{B(x_0,r)} u(x)\,dx.$$

Since $u \le A$, equality of the average with A forces $u \equiv A$ on $B(x_0, r)$; otherwise the average would be strictly smaller than A.

Equivalently, the set $E = \{x \in \Omega : u(x) = A\}$ is nonempty and open. It is also closed in Ω by continuity. Connectedness therefore implies $E = \Omega$. $\square$

Theorem 3.4 (Boundary maximum principle)**.** *Let $\Omega \subset \mathbb{R}^m$ be bounded and open, and suppose $u \in C(\overline{\Omega})$ is harmonic in Ω. Then*

$$\max_{\overline{\Omega}} u = \max_{\partial\Omega} u, \qquad \min_{\overline{\Omega}} u = \min_{\partial\Omega} u.$$

Proof. Compactness of $\overline{\Omega}$ guarantees that the extrema are attained. If a maximum occurs in the interior of a connected component, the strong maximum principle makes u constant on that component, so the same value occurs on its boundary. Apply the same argument to $-u$ for the minimum. □

Theorem 3.5 (Uniqueness for the Dirichlet problem)**.** *Let Ω be bounded and let $f \in C(\partial\Omega)$. The problem*

$$\begin{cases} \Delta u = 0 & \text{in } \Omega, \\ u = f & \text{on } \partial\Omega, \end{cases} \qquad u \in C(\overline{\Omega}),$$

has at most one solution.

Proof. If u_1 and u_2 are two solutions, then $v = u_1 - u_2$ satisfies

$$\Delta v = 0 \quad \text{in } \Omega, \qquad v = 0 \quad \text{on } \partial\Omega.$$

The boundary maximum principle applied to v and $-v$ gives $\max_{\overline{\Omega}} v = \min_{\overline{\Omega}} v = 0$. Hence $v \equiv 0$. □

3.2 The fundamental solution of the Laplacian

Assume $m \geq 3$, and write $\omega_m = |\mathbb{S}^{m-1}|$. Define the signed fundamental solution

$$\Gamma(x) = \frac{1}{(2-m)\omega_m} \frac{1}{|x|^{m-2}}, \qquad x \neq 0. \tag{3.2}$$

For every $\varphi \in C_c^2(\mathbb{R}^m)$,

$$\int_{\mathbb{R}^m} \Gamma(x) \Delta\varphi(x)\, dx = \varphi(0). \tag{3.3}$$

Thus

$$\Delta\Gamma = \delta_0$$

in the sense of distributions. Translation gives

$$\int_{\mathbb{R}^m} \Delta\varphi(y)\Gamma(x-y)\, dy = \varphi(x). \tag{3.4}$$

Consequently, if

$$v(x) = \int_{\mathbb{R}^m} \varphi(y)\Gamma(x-y)\, dy$$

is twice differentiable, then moving the derivatives onto φ gives

$$\Delta v(x) = \int_{\mathbb{R}^m} \Delta\varphi(x-y)\Gamma(y)\,dy = \varphi(x).$$

The phrase "if v is twice differentiable" is important: the second derivatives of Γ are too singular to be ordinary locally integrable functions.

Indeed,

$$\partial_i |z|^{2-m} = (2-m)\frac{z_i}{|z|^m},$$

and hence

$$|\partial_i \Gamma(z)| \le C|z|^{1-m} \in L^1_{\mathrm{loc}}.$$

On the other hand,

$$|\partial_{ij}\Gamma(z)| \le C|z|^{-m}, \qquad |z|^{-m} \notin L^1_{\mathrm{loc}}.$$

This is a borderline failure: $|z|^{-m+\varepsilon}$ is locally integrable for every $\varepsilon > 0$. To recover second derivatives, we must exploit cancellation rather than absolute integrability.

3.3 Principal values and cancellation

The one-dimensional kernel $1/x$ is the model example. For $\varphi \in C^1([-1,1])$, define

$$\mathrm{p.v.}\int_{-1}^{1} \frac{\varphi(x)}{x}\,dx = \lim_{\varepsilon \downarrow 0}\left(\int_{-1}^{-\varepsilon} \frac{\varphi(x)}{x}\,dx + \int_{\varepsilon}^{1} \frac{\varphi(x)}{x}\,dx\right).$$

Write

$$\varphi(x) = \varphi(0) + xR(x), \qquad R(x) = \frac{\varphi(x)-\varphi(0)}{x},$$

where R extends continuously to $R(0) = \varphi'(0)$. The singular constant term cancels because

$$\int_{-1}^{-\varepsilon} \frac{dx}{x} + \int_{\varepsilon}^{1} \frac{dx}{x} = 0,$$

while R is integrable. Therefore the principal value is well defined and equals $\int_{-1}^{1} R(x)\,dx$.

The same mechanism works in $\mathbb{R}^m$. Suppose $K\colon \mathbb{R}^m \setminus \{0\} \to \mathbb{R}$ is homogeneous of degree $-m$:

$$K(rz) = r^{-m}K(z), \qquad r > 0. \tag{3.5}$$

Writing $z = r\omega$, with $|\omega| = 1$, gives

$$K(z) = \frac{K(\omega)}{|z|^m},$$

which is not locally integrable at the origin. If $f(r\omega) = f(0) + rR(r,\omega)$, then the truncated integral has the form

$$\int_{\varepsilon<|z|<b} f(z)K(z)\,dz = f(0)\int_{\varepsilon}^{b} \frac{dr}{r}\int_{\mathbb{S}^{m-1}} K(\omega)\,d\omega$$

$$+\int_{\varepsilon}^{b}\int_{\mathbb{S}^{m-1}} R(r,\omega)K(\omega)\,d\omega\,dr.$$

The second term is locally integrable. The logarithmic first term can cancel for arbitrary f only if

$$\int_{\mathbb{S}^{m-1}} K(\omega)\,d\omega = 0. \tag{3.6}$$

Under this cancellation condition, the integral of K over every annulus vanishes:

$$\int_{a<|z|<b} K(z)\,dz = \big(\log b - \log a\big)\int_{\mathbb{S}^{m-1}} K(\omega)\,d\omega = 0.$$

Conversely, vanishing on every annulus implies (3.6). We may therefore subtract the value at the singularity:

$$\text{p. v.}\int f(z)K(z)\,dz = \lim_{b\to\infty}\int_{|z|<b}\big(f(z)-f(0)\big)K(z)\,dz, \tag{3.7}$$

whenever the large-scale limit is meaningful. Near zero, Hölder continuity is already sufficient: if

$$|f(z)-f(0)| \le C|z|^{\alpha}, \qquad 0<\alpha<1,$$

then the integrand in (3.7) is bounded in absolute value by $C|z|^{\alpha-m}$, which is locally integrable.

3.4 Hölder spaces and singular integral operators

Let $0<\alpha<1$. For a bounded domain Ω, define

$$[f]_{C^{\alpha}(\overline{\Omega})} = \sup_{\substack{x,y\in\overline{\Omega}\\ x\ne y}} \frac{|f(x)-f(y)|}{|x-y|^{\alpha}}, \qquad \|f\|_{C^{\alpha}(\overline{\Omega})} = \|f\|_{C(\overline{\Omega})} + [f]_{C^{\alpha}(\overline{\Omega})}.$$

The space $C^{\alpha}(\overline{\Omega})$, equipped with this norm, is a Banach space. More generally,

$$\|f\|_{C^{k,\alpha}(\overline{\Omega})} = \|f\|_{C^{k}(\overline{\Omega})} + \max_{|\beta|=k}[\partial^{\beta} f]_{C^{\alpha}(\overline{\Omega})}.$$

Definition 3.6. A function $K\colon \mathbb{R}^m\setminus\{0\}\to\mathbb{R}$ is a *singular integral kernel* if

1. $K(rz) = r^{-m}K(z)$ for every $r>0$, and $\int_{a<|z|<b} K(z)\,dz = 0$ for every $0<a<b$;
2. $K\in C^{1}(\mathbb{R}^m\setminus\{0\})$.

For $f\in C_c^{\alpha}(\mathbb{R}^m)$, define

$$Tf(x) = \text{p. v.}\int_{\mathbb{R}^m} f(x-y)K(y)\,dy \tag{3.8}$$

$$= \lim_{b\to\infty}\int_{|y|<b}\big(f(x-y)-f(x)\big)K(y)\,dy. \tag{3.9}$$

The subtraction makes the singularity integrable, while compact support makes the large-scale truncation unambiguous. The preceding estimates also give $Tf\in L^{\infty}(\mathbb{R}^m)$. It remains to control the Hölder seminorm.

Theorem 3.7 (Hölder continuity of singular integrals). *If K is a singular integral kernel and $f \in C_c^\alpha(\mathbb{R}^m)$, then*

$$Tf \in C^\alpha(\mathbb{R}^m).$$

Proof. The expression (3.9) is well defined by the preceding cancellation argument. Fix a displacement y, put $r = |y|$, and split $Tf = w_1 + w_2$, where

$$w_1(x) = \int_{|z|<3r} \big(f(x-z) - f(x)\big)K(z)\,dz,$$
$$w_2(x) = \lim_{b\to\infty} \int_{3r<|z|<b} \big(f(x-z) - f(x)\big)K(z)\,dz.$$

Since $|K(z)| \le C|z|^{-m}$,

$$|w_1(x)| \le C[f]_{C^\alpha} \int_{|z|<3r} |z|^{\alpha-m}\,dz \le C[f]_{C^\alpha} r^\alpha.$$

The same bound holds at $x+y$, so

$$|w_1(x+y) - w_1(x)| \le C[f]_{C^\alpha} r^\alpha. \tag{3.10}$$

To compare the far terms, use annular cancellation to replace $f(x+y)$ by $f(x)$ in $w_2(x+y)$, and then make the change of variables $z \mapsto z+y$. On the common part of the two truncation regions, the resulting difference is

$$\big(f(x-z) - f(x)\big)\big(K(z+y) - K(z)\big).$$

The discrepancy between the two outer boundaries gives an error $E_1(b)$ supported where $b - r < |z| < b + r$. Hence

$$|E_1(b)| \le C\|f\|_{L^\infty} \frac{(b+r)^m - (b-r)^m}{b^m} \longrightarrow 0$$

as $b \longrightarrow \infty$.

The discrepancy between the inner boundaries gives an error E_2 supported in the annulus $2r < |z| < 4r$. In that region $|z| \simeq |z+y| \simeq r$, and therefore

$$|E_2| \le C[f]_{C^\alpha} r^\alpha.$$

Finally, homogeneity and C^1-regularity imply $|\partial_i K(z)| \le C|z|^{-m-1}$. The ordinary mean-value theorem, together with $|z| > 3r$, gives

$$|K(z+y) - K(z)| \le C\frac{r}{|z|^{m+1}}.$$

Thus the main term is bounded by

$$\begin{aligned} C[f]_{C^\alpha} \int_{|z|>3r} |z|^\alpha \frac{r}{|z|^{m+1}}\,dz &\le C[f]_{C^\alpha} r \int_{3r}^\infty \rho^{\alpha-2}\,d\rho \\ &\le C[f]_{C^\alpha} r^\alpha, \end{aligned}$$

because $0 < \alpha < 1$. Combining this estimate with (3.10) and the two boundary errors proves $|Tf(x+y) - Tf(x)| \le C|y|^\alpha$. □

3.5 Regularity of the signed Newtonian potential

Theorem 3.8 (Schauder regularity for the signed Newtonian potential)**.** *Let $m \geq 3$, $0 < \alpha < 1$, and $f \in C_c^\alpha(\mathbb{R}^m)$. If*

$$u = f * \Gamma,$$

with Γ given by (3.2), then

$$u \in C^{2,\alpha}(\mathbb{R}^m) \qquad \textit{and} \qquad \Delta u = f.$$

Proof. The distributional identity $\Delta\Gamma = \delta_0$ gives $\Delta u = f$. Away from the origin, each kernel $\partial_{ij}\Gamma$ is homogeneous of degree $-m$, is C^1, and has zero spherical mean. Distributionally, $\partial_{ij}\Gamma$ is the corresponding principal-value kernel together with a constant multiple of δ_0. Therefore each $\partial_{ij}u$ is the sum of a singular integral of f and a constant multiple of f. The preceding theorem yields $\partial_{ij}u \in C^\alpha$, and hence $u \in C^{2,\alpha}$. □

3.6 Problems

Problem 3.1. Consider the equation

$$\partial_{xx}u + 2\partial_{xt}u + \partial_{tt}u = 0.$$

(a) Find its characteristic equation and classify the equation.
(b) Derive the general solution.
(c) Solve the Cauchy problem

$$u(x,0) = f(x), \qquad \partial_t u(x,0) = g(x),$$

and state the local uniqueness conclusion when the data are analytic.

Solution. The characteristic equation is

$$(\partial_x\phi)^2 + 2(\partial_x\phi)(\partial_t\phi) + (\partial_t\phi)^2 = 0,$$

or

$$\partial_x\phi = -\partial_t\phi$$

with a double root. Thus the discriminant vanishes and the equation is parabolic.

The repeated characteristic leads to a solution of the form

$$u(x,t) = F(x-t) + tG(x-t).$$

Indeed,

$$\begin{aligned} \partial_x u &= F' + tG', & \partial_{xx}u &= F'' + tG'', \\ \partial_t u &= -F' + G - tG', & \partial_{tt}u &= F'' - 2G' + tG'', \\ 2\partial_{tx}u &= -2F'' + 2G' - 2tG''. \end{aligned}$$

Consequently,

$$\partial_{xx}u + 2\partial_{tx}u + \partial_{tt}u = 0.$$

At $t = 0$, the prescribed data give

$$F(x) = f(x)$$

and

$$G(x) - F'(x) = g(x), \qquad \text{so} \qquad G(x) = f'(x) + g(x).$$

Therefore

$$\boxed{u(x,t) = f(x-t) + t\big(f'(x-t) + g(x-t)\big)}.$$

For analytic f and g, Holmgren's theorem gives local uniqueness near the initial curve $t = 0$. □

Problem 3.2. Consider the Cauchy problem

$$\partial_{tt}u = \partial_{xx}u + u, \qquad u(x,0) = \mathrm{e}^x, \qquad \partial_t u(x,0) = 0.$$

Determine the power series in t obtained from the Cauchy–Kovalevskaya recursion.

Solution. For every $k \in \mathbb{N}_0$,

$$\partial_x^k u(x,0) = \mathrm{e}^x.$$

Using the equation repeatedly gives

$$\begin{aligned}
\partial_{tt}u(x,0) &= \partial_{xx}u(x,0) + u(x,0) = 2\mathrm{e}^x,\\
\partial_t^3 u(x,0) &= \partial_{xx}\partial_t u(x,0) + \partial_t u(x,0) = 0,\\
\partial_t^4 u(x,0) &= \partial_{xx}\partial_t^2 u(x,0) + \partial_t^2 u(x,0) = 4\mathrm{e}^x,\\
\partial_t^5 u(x,0) &= 0,\\
\partial_t^6 u(x,0) &= \partial_{xx}\partial_t^4 u(x,0) + \partial_t^4 u(x,0) = 8\mathrm{e}^x.
\end{aligned}$$

Thus, by induction,

$$\partial_t^{2k} u(x,0) = 2^k \mathrm{e}^x, \qquad \partial_t^{2k+1} u(x,0) = 0.$$

The Cauchy–Kovalevskaya power series is therefore

$$\boxed{u(x,t) = \sum_{k=0}^{\infty} \frac{2^k \mathrm{e}^x}{(2k)!}\, t^{2k}}.$$

It converges in a neighborhood of $t = 0$. □

Problem 3.3. Let a second-order equation in two variables have characteristic equation

$$a(\partial_{x_1}\phi)^2 + 2b(\partial_{x_1}\phi)(\partial_{x_2}\phi) + c(\partial_{x_2}\phi)^2 = 0.$$

Classify the equation in terms of

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

and prove that its type is preserved under every invertible linear change of variables $y = Ax$.

Solution. The characteristic equation can be written as

$$\left\langle M\nabla_x\phi, \nabla_x\phi\right\rangle = 0, \qquad \det M = ac - b^2.$$

Hence the equation is hyperbolic if $\det M < 0$, parabolic if $\det M = 0$, and elliptic if $\det M > 0$.

Let $y = Ax$, where $\det A \neq 0$, and write $\widetilde{\phi}(y) = \phi(A^{-1}y)$. The chain rule gives

$$\nabla_x\phi = A^{\mathsf{T}}\nabla_y\widetilde{\phi}.$$

Thus, in the y variables, the characteristic equation becomes

$$\begin{aligned} 0 &= \left\langle MA^{\mathsf{T}}\nabla_y\widetilde{\phi}, A^{\mathsf{T}}\nabla_y\widetilde{\phi}\right\rangle \\ &= \left\langle AMA^{\mathsf{T}}\nabla_y\widetilde{\phi}, \nabla_y\widetilde{\phi}\right\rangle. \end{aligned}$$

If $C = AMA^{\mathsf{T}}$ is the transformed principal matrix, then

$$\det C = \det(AMA^{\mathsf{T}}) = (\det A)^2 \det M.$$

Since $(\det A)^2 > 0$, the signs of $\det C$ and $\det M$ agree. The classification is therefore preserved. □

Problem 3.4. Let $m \geq 3$, let $\phi \in C_c^\infty(\mathbb{R}^m)$, and write $\omega_m = \left|\mathbb{S}^{m-1}\right|$. Prove that

$$\int_{\mathbb{R}^m} \Delta\phi(x)\frac{1}{|x|^{m-2}}\,\mathrm{d}x = -(m-2)\omega_m\phi(0).$$

Solution. For $\varepsilon > 0$, split the integral as

$$\int_{\mathbb{R}^m} \Delta\phi(x)\frac{1}{|x|^{m-2}}\,\mathrm{d}x = I_\varepsilon + J_\varepsilon,$$

where

$$I_\varepsilon = \int_{B_\varepsilon} \Delta\phi(x)\frac{1}{|x|^{m-2}}\,\mathrm{d}x, \qquad J_\varepsilon = \int_{\mathbb{R}^m\setminus B_\varepsilon} \Delta\phi(x)\frac{1}{|x|^{m-2}}\,\mathrm{d}x.$$

The first term satisfies

$$\begin{aligned} |I_\varepsilon| &\leq \left\|\Delta\phi\right\|_{L^\infty}\int_{B_\varepsilon}\frac{1}{|x|^{m-2}}\,\mathrm{d}x \\ &= \omega_m\left\|\Delta\phi\right\|_{L^\infty}\int_0^\varepsilon r\,\mathrm{d}r = \frac{\omega_m}{2}\left\|\Delta\phi\right\|_{L^\infty}\varepsilon^2, \end{aligned}$$

so $I_\varepsilon \to 0$.

Set $\Gamma(x) = |x|^{2-m}$. Since $\Delta\Gamma = 0$ away from the origin, Green's identity on $\mathbb{R}^m \setminus B_\varepsilon$ gives

$$J_\varepsilon = \int_{\partial B_\varepsilon}\left(\Gamma\partial_\nu\phi - \phi\partial_\nu\Gamma\right)\,\mathrm{d}S_x,$$

where $\nu = -x/|x|$ is the outward unit normal for the punctured domain. Now

$$\nabla\Gamma(x) = -(m-2)\frac{x}{|x|^m}, \qquad \partial_\nu\Gamma = (m-2)\varepsilon^{1-m} \quad \text{on } \partial B_\varepsilon.$$

The first boundary term tends to zero, because

$$\left|\int_{\partial B_\varepsilon} \Gamma \partial_\nu \phi \, \mathrm{d}S_x\right| \leq \left\|\nabla \phi\right\|_{L^\infty} \varepsilon^{2-m} \omega_m \varepsilon^{m-1} = \omega_m \left\|\nabla \phi\right\|_{L^\infty} \varepsilon.$$

For the second boundary term,

$$-(m-2)\varepsilon^{1-m} \int_{\partial B_\varepsilon} \phi(x) \, \mathrm{d}S_x = -(m-2)\omega_m \left(\frac{1}{\omega_m \varepsilon^{m-1}} \int_{\partial B_\varepsilon} \phi(x) \, \mathrm{d}S_x\right).$$

The normalized spherical average in parentheses converges to $\phi(0)$ as $\varepsilon \downarrow 0$. Therefore

$$J_\varepsilon \longrightarrow -(m-2)\omega_m \phi(0).$$

Combining this with $I_\varepsilon \to 0$ proves the identity. □

Chapter 4

Green Functions and Distributions

4.1 Green's representation for harmonic functions

Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with smooth boundary, and let $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ satisfy

$$\Delta u = 0 \qquad \text{in } \Omega.$$

For $m \geq 3$, write

$$N(z) = \frac{1}{(m-2)|\mathbb{S}^{m-1}|} \frac{1}{|z|^{m-2}}.$$

Thus $N = -\Gamma$, where Γ was the signed fundamental solution in the preceding chapter. The two symbols keep the sign conventions explicit: $\Delta\Gamma = \delta_0$ and $-\Delta N = \delta_0$. With this normalization, $-\Delta N = \delta_0$ in the sense of distributions. Thus, for a fixed $x \in \Omega$, Green's second identity applied to $u(y)$ and $N(x-y)$, first on $\Omega \setminus \overline{B_\varepsilon(x)}$ and then in the limit $\varepsilon \downarrow 0$, gives

$$u(x) = \int_{\partial\Omega} N(x-y)\, \partial_\nu u(y)\, dS_y - \int_{\partial\Omega} u(y)\, \partial_{\nu_y} N(x-y)\, dS_y. \tag{4.1}$$

Here ν_y is the outward unit normal at $y \in \partial\Omega$.

Formula (4.1) is tempting because it represents u in terms of both u and $\partial_\nu u$ on the boundary. It is not yet the natural formula for the Dirichlet problem: the boundary value $u|_{\partial\Omega}$ already determines a harmonic function uniquely, so the normal derivative should not have to be prescribed independently.

To remove that term, consider

$$G(x,y) = -N(x-y) + K(x,y),$$

where, for each fixed $x \in \Omega$, the correction $K(x,\cdot)$ is chosen to satisfy

$$\Delta_y K(x,y) = 0 \quad \text{in } \Omega, \qquad K(x,y) = N(x-y) \quad \text{on } \partial\Omega.$$

Then

$$\Delta_y G(x,y) = \delta_x(y), \qquad G(x,y) = 0 \quad \text{for } y \in \partial\Omega.$$

Definition 4.1 (Green function)**.** For the sign convention above, a function G with these two properties is called the Green function of Ω for the Laplacian.

Green's identity now yields

$$u(x) = \int_{\partial\Omega} u(y)\,\partial_{\nu_y} G(x,y)\,dS_y - \int_{\partial\Omega} G(x,y)\,\partial_\nu u(y)\,dS_y.$$

The second integral vanishes because $G = 0$ on $\partial\Omega$. Hence

$$\boxed{u(x) = \int_{\partial\Omega} u(y)\,\partial_{\nu_y} G(x,y)\,dS_y.} \tag{4.2}$$

This is the desired representation in terms of Dirichlet data alone. Of course, constructing G requires solving the auxiliary harmonic problem for K.

4.1.1 The Newtonian potential

For a function f for which the integral is defined, set

$$v(x) = \int_{\mathbb{R}^m} f(y)N(x-y)\,dy. \tag{4.3}$$

The distributional identity $-\Delta N = \delta_0$ gives

$$\Delta v = -f.$$

The potential improves regularity by two derivatives in the Hölder scale: in the usual local form, if $f \in C^{0,\alpha}_{\mathrm{loc}}$ and the potential is well defined, then

$$v \in C^{2,\alpha}_{\mathrm{loc}}.$$

This regularity conclusion can fail if f is only continuous; there are continuous right-hand sides whose Newtonian potentials are not C^2. This motivates a broader notion of solution.

4.2 Why the Cauchy problem for the Laplacian is ill posed

Consider the Cauchy data on the line $y = 0$:

$$\partial_x^2 u + \partial_y^2 u = 0, \qquad u(x,0) = \delta\sin(mx), \qquad \partial_y u(x,0) = 0. \tag{4.4}$$

The solution can be written explicitly as

$$u(x,y) = \delta\sin(mx)\cosh(my).$$

Taking $\delta = 1/m$, the prescribed value on $y = 0$ tends uniformly to zero as $m \to \infty$, while on any fixed line $y \neq 0$,

$$\sup_x |u(x,y)| = \frac{\cosh(m|y|)}{m} \longrightarrow \infty.$$

Thus arbitrarily small Cauchy data in the uniform norm can produce arbitrarily large interior values. The Cauchy problem for the Laplacian does not have continuous dependence on the data and is therefore ill posed. This contrasts with the Dirichlet problem, which is well posed under the hypotheses studied earlier.

4.3 Test functions and weak equations

Suppose $f \in L^1_{\mathrm{loc}}(\Omega)$. There are two equivalent ways to say that $f = 0$:

$$f = 0 \quad \text{almost everywhere,} \qquad \text{or} \qquad \int_\Omega f\phi\,dx = 0 \quad \text{for every } \phi \in C_c^\infty(\Omega).$$

This observation suggests testing an equation rather than differentiating its putative solution. If a classical function u satisfies $\Delta u = f$, then integration by parts gives

$$\int_\Omega u\,\Delta\phi\,dx = \int_\Omega f\phi\,dx \qquad \text{for every } \phi \in C_c^\infty(\Omega).$$

Definition 4.2 (Distributional solution)**.** A locally integrable function u is a distributional solution of $\Delta u = f$ if it satisfies the displayed test-function identity above. No classical differentiability of u is required.

The notation

$$\mathcal{D}(\Omega) = C_c^\infty(\Omega)$$

will be used for the space of test functions. Thus $\phi \in \mathcal{D}(\Omega)$ means that ϕ is smooth and that $\operatorname{supp}\phi \Subset \Omega$.

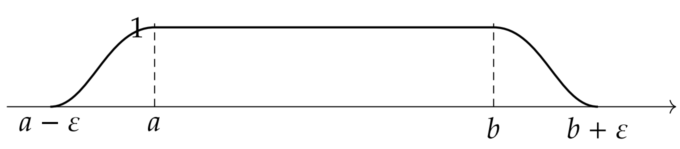
Taylor's theorem with integral remainder will be used repeatedly in estimates for test functions. If $x, x_0 \in (a, b)$ and the segment joining them lies in (a, b), then

$$\begin{aligned}\phi(x) = {} & \sum_{j=0}^{k} \frac{\phi^{(j)}(x_0)}{j!}(x - x_0)^j \\ & + \frac{(x - x_0)^{k+1}}{k!} \int_0^1 (1 - s)^k \phi^{(k+1)}\big(x_0 + s(x - x_0)\big)\,ds.\end{aligned} \tag{4.5}$$

Smooth cutoff functions make this space rich. For example, given $a < b$ and $\varepsilon > 0$, one can construct $\phi \in C_c^\infty(\mathbb{R})$ such that

$$0 \le \phi \le 1, \qquad \phi = 1 \text{ on } [a, b], \qquad \operatorname{supp}\phi \subset [a - \varepsilon, b + \varepsilon].$$

One construction begins with a nonnegative bump χ of integral one, sets $\chi_\varepsilon(x) = \varepsilon^{-1}\chi(x/\varepsilon)$, and convolves a suitable interval's characteristic function with χ_ε.



4.3.1 Partitions of unity

If a compact set K is covered by finitely many open sets,

$$K \subset \bigcup_{i=1}^{N} O_i,$$

then there exist $\phi_i \in C_c^\infty(O_i)$ such that

$$\sum_{i=1}^{N} \phi_i = 1 \qquad \text{on } K.$$

More generally, for a locally finite open cover $\{O_i\}_{i=1}^\infty$, one may choose $\phi_i \in C_c^\infty(O_i)$ so that the sum is locally finite and

$$\sum_{i=1}^{\infty} \phi_i(x) = 1 \qquad \text{for } x \in \bigcup_{i=1}^{\infty} O_i.$$

The notation $A \Subset S$ means that $\overline{A}$ is a compact subset of S.

4.4 Distributions

The correct notion of convergence on $\mathcal{D}(a,b)$ is the following. A sequence ϕ_j converges to ϕ in $\mathcal{D}(a,b)$ if there is a compact set $K \Subset (a,b)$ containing every $\operatorname{supp}\phi_j$ and $\operatorname{supp}\phi$, and if

$$\|\phi_j - \phi\|_{C^N(K)} \longrightarrow 0 \qquad \text{for every } N \geq 0,$$

where

$$\|\psi\|_{C^N(K)} = \sup_{x\in K} \sum_{k=0}^{N} |\psi^{(k)}(x)|.$$

Definition 4.3 (Distribution)**.** A distribution on (a,b) is a linear functional

$$T : \mathcal{D}(a,b) \longrightarrow \mathbb{R}$$

that is continuous for this convergence. The space of distributions is denoted

$$\mathcal{D}'(a,b) = \left(C_c^\infty(a,b)\right)^*.$$

The action of T on ϕ is written $\langle T, \phi\rangle$.

Continuity is equivalent to the following local boundedness statement: for every compact $K \Subset (a,b)$, there are a constant C_K and an integer $N(K)$ such that

$$|\langle T, \phi\rangle| \leq C_K \|\phi\|_{C^{N(K)}(K)}$$

whenever $\phi \in C_c^\infty(a,b)$ and $\operatorname{supp}\phi \subset K$.

4.4.1 Regular and singular examples

Every $u \in L^1_{\mathrm{loc}}(a,b)$ defines a distribution

$$\langle T_u, \phi\rangle = \int_a^b u(x)\phi(x)\,dx.$$

Indeed, for $\operatorname{supp}\phi \subset K \Subset (a,b)$,

$$\left|\int_a^b u\phi\,dx\right| \le \left(\int_K |u|\,dx\right)\|\phi\|_{C^0(K)}.$$

Such distributions are called regular distributions.

The Dirac mass at x_0 is defined by

$$\langle \delta_{x_0}, \phi\rangle = \phi(x_0).$$

It is a distribution, but it is not induced by a locally integrable function.

Another singular example is the principal value of $1/x$:

$$\begin{aligned}\left\langle \mathrm{p.\,v.}\,\frac{1}{x}, \phi\right\rangle &:= \lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \frac{\phi(x)}{x}\,dx \\ &= \lim_{\varepsilon\downarrow 0}\int_\varepsilon^\infty \frac{\phi(x)-\phi(-x)}{x}\,dx.\end{aligned}$$

If $\operatorname{supp}\phi \subset [-R,R]$, the mean value theorem gives

$$\left|\left\langle \mathrm{p.\,v.}\,\frac{1}{x}, \phi\right\rangle\right| \le 2R \sup_{|x|\le R}|\phi'(x)|,$$

so the principal value is a distribution.

4.5 Differentiation of distributions

Definition 4.4 (Distributional derivative)**.** For $T \in \mathcal{D}'(a,b)$, define its derivative $DT = T'$ by

$$\langle T', \phi\rangle = -\langle T, \phi'\rangle.$$

This extends ordinary differentiation because integration by parts gives the same formula for every differentiable regular distribution.

Every differential operator defined on test functions therefore extends to distributions by duality.

4.5.1 The Heaviside function

Let

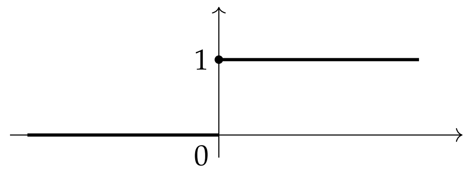
$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases}$$

Its classical derivative vanishes almost everywhere, but its distributional derivative records the jump:

$$\langle DH, \phi\rangle = -\int_0^\infty \phi'(x)\,dx = \phi(0) = \langle \delta_0, \phi\rangle.$$

Thus

$$\boxed{DH = \delta_0.} \tag{4.6}$$



Since $\log|x| \in L^1_{\mathrm{loc}}(\mathbb{R})$, it also defines a distribution, and

$$D\log|x| = \mathrm{p.\,v.}\,\frac{1}{x}.$$

If $f \in L^1_{\mathrm{loc}}(a,b)$ and

$$F(x) = \int_{x_0}^{x} f(t)\,dt,$$

then F is locally absolutely continuous and

$$DF = f$$

as distributions. The distributional statement avoids confusing an almost-everywhere derivative with complete recovery of a function. For example, the Cantor function is nonconstant and increasing even though its classical derivative is zero almost everywhere.

4.5.2 Distributions with zero derivative

Lemma 4.5. *If $T \in \mathcal{D}'(\mathbb{R})$ and $T' = 0$, then T is a constant distribution: there is a constant C such that*

$$\langle T, \phi\rangle = C\int_{\mathbb{R}} \phi(x)\,dx \qquad \textit{for every } \phi \in \mathcal{D}(\mathbb{R}).$$

Proof. First take $\psi \in \mathcal{D}(\mathbb{R})$ with $\int \psi = 0$. Then

$$\Phi(x) = \int_{-\infty}^{x} \psi(t)\,dt$$

belongs to $\mathcal{D}(\mathbb{R})$, and

$$\langle T, \psi\rangle = \langle T, \Phi'\rangle = -\langle T', \Phi\rangle = 0.$$

For a general ψ, choose $\chi \in \mathcal{D}(\mathbb{R})$ with $\int \chi = 1$, set $c_\psi = \int \psi$, and apply the preceding argument to $\psi - c_\psi\chi$. It follows that

$$\langle T, \psi\rangle = c_\psi\langle T, \chi\rangle,$$

which is the desired formula with $C = \langle T, \chi\rangle$. □

Lemma 4.6 (Antiderivatives in $\mathcal{D}'$)**.** *For every $S \in \mathcal{D}'(\mathbb{R})$, there exists $T \in \mathcal{D}'(\mathbb{R})$ such that*

$$DT = S.$$

Proof. Fix $\chi \in \mathcal{D}(\mathbb{R})$ with $\int \chi = 1$. For $\phi \in \mathcal{D}(\mathbb{R})$, set $c_\phi = \int \phi$ and define

$$M\phi(x) = \int_{-\infty}^{x} \big(\phi(t) - c_\phi \chi(t)\big)\, dt.$$

The integrand has integral zero, so $M\phi$ is again a compactly supported smooth function. Define

$$\langle T, \phi \rangle = -\langle S, M\phi \rangle.$$

The map M is continuous on test functions, so T is a distribution. Moreover, $c_{\phi'} = 0$ and $M(\phi') = \phi$. Therefore

$$\langle DT, \phi \rangle = -\langle T, \phi' \rangle = \langle S, M(\phi') \rangle = \langle S, \phi \rangle.$$

□

4.6 Multiplication and elementary distributional equations

If $f \in C^\infty(\mathbb{R})$ and $T \in \mathcal{D}'(\mathbb{R})$, define the product fT by

$$\langle fT, \phi \rangle = \langle T, f\phi \rangle.$$

This is well defined because $f\phi \in C_c^\infty(\mathbb{R})$. The ordinary product rule remains valid:

$$D(fT) = f'T + f\, DT. \tag{4.7}$$

Two useful consequences are

$$(x - x_0)T = 0 \quad \Longrightarrow \quad T = c\, \delta_{x_0},$$

and, more generally,

$$(x - x_0)^N T = 0 \quad \Longrightarrow \quad T = \sum_{j=0}^{N-1} c_j\, \delta_{x_0}^{(j)}.$$

As an application, consider first the classical equation

$$xy'(x) + y(x) = 0.$$

Since $(xy)' = 0$, its classical solutions away from $x = 0$ have the form $y = c/x$. The distributional equation is

$$xT' + T = 0, \qquad \text{or equivalently} \qquad D(xT) = 0.$$

The zero-derivative lemma gives $xT = c$, where the right-hand side is a constant distribution. On the other hand,

$$x\, \mathrm{p.v.}\, \frac{1}{x} = 1,$$

because

$$\left\langle x\, \mathrm{p.v.}\, \frac{1}{x}, \phi \right\rangle = \int_{\mathbb{R}} \phi(x)\, dx.$$

It follows that $x\big(T - c\,\mathrm{p.v.}(1/x)\big) = 0$, and the first consequence above gives the complete family

$$\boxed{T = c_1\delta_0 + c_2\,\mathrm{p.v.}\,\frac{1}{x}.} \tag{4.8}$$

Likewise, if a polynomial P satisfies $P(x)T = 0$, one factors P and applies the preceding point-support statements. In particular, if the distinct real roots are a_j, with multiplicities m_j, then every such distribution is a finite linear combination of

$$\delta_{a_j}, \delta'_{a_j}, \dots, \delta_{a_j}^{(m_j-1)}.$$

4.7 Tempered distributions

Ordinary distributions are controlled on each fixed compact set, but the constants and the number of derivatives may depend on that compact set. Tempered distributions impose a single global condition and therefore behave well at infinity.

For $p \in \mathbb{N}$, introduce the Schwartz seminorm

$$\mathcal{N}_p(\phi) = \sup_{x\in\mathbb{R}^m} \sum_{|\alpha|+|\beta|\le p} \left| x^\alpha \partial^\beta \phi(x) \right|.$$

The Schwartz space $\mathcal{S}(\mathbb{R}^m)$ consists of the smooth functions for which every $\mathcal{N}_p$ is finite.

Definition 4.7 (Tempered distribution)**.** A distribution T is tempered if it extends continuously to $\mathcal{S}(\mathbb{R}^m)$. Equivalently, there are constants $C > 0$ and $p \in \mathbb{N}$, independent of the support of ϕ, such that

$$|\langle T, \phi\rangle| \le C\mathcal{N}_p(\phi)$$

for every $\phi \in C_c^\infty(\mathbb{R}^m)$. The space of tempered distributions is denoted $\mathcal{S}'(\mathbb{R}^m)$.

4.8 Problems

Problem 4.1. Let

$$\mathbb{R}^d_+ := \left\{ x = (x', x_d) \in \mathbb{R}^d : x_d > 0 \right\}.$$

Suppose that

$$u \in C^2(\mathbb{R}^d_+) \cap C(\overline{\mathbb{R}^d_+}), \qquad \Delta u = 0 \quad \text{in } \mathbb{R}^d_+,$$

and that $u(x) \to 0$ as $|x| \to \infty$. Prove the maximum principle on the upper half-space:

$$\sup_{\overline{\mathbb{R}^d_+}} u = \sup_{x'\in\mathbb{R}^{d-1}} u(x', 0).$$

Solution. For $R > 0$, set

$$S_R := \mathbb{R}^d_+ \cap B(0, R).$$

The maximum principle on the bounded domain S_R gives

$$\sup_{\overline{S_R}} u = \sup_{\partial S_R} u = \max\left\{ \sup_{\substack{x_d=0\\ |x'|\le R}} u(x',0),\ \sup_{\substack{x\in\overline{\mathbb{R}^d_+}\\ |x|=R}} u(x) \right\}.$$

As $R \to \infty$, the second term tends to 0. Moreover, $u(x',0) \to 0$ as $|x'| \to \infty$, so $\sup_{x'} u(x',0) \ge 0$. Hence

$$\sup_{\overline{\mathbb{R}^d_+}} u \le \sup_{x'\in\mathbb{R}^{d-1}} u(x',0).$$

The reverse inequality is immediate because $\partial\mathbb{R}^d_+ \subset \overline{\mathbb{R}^d_+}$. □

Problem 4.2. Let $m \ge 3$, and let N be the Newtonian fundamental solution normalized by

$$\Delta N = \delta_0, \qquad |N(x)| \lesssim |x|^{2-m}.$$

Suppose that f is smooth and, for some $A > 2$,

$$\left|f(y)\right| \lesssim \left|y\right|^{-A} \qquad \text{for large } \left|y\right|.$$

(a) Show that the Newtonian potential

$$u(x) := \int_{\mathbb{R}^m} f(y)N(x-y)\,\mathrm{d}y$$

is well defined and smooth. For the estimate recorded in the calculation,

$$|u(x)| \lesssim |x|^{2-A} + |x|,$$

determine the conditions on B under which its weighted right-hand side tends to zero as $|x| \to \infty$, and as $|x| \to 0$.

(b) Let

$$T = \operatorname{diag}(1,\dots,1,-1), \qquad \widetilde{f} := f \circ T,$$

and let $v := N * \widetilde{f}$. Prove that

$$v(x) = u(Tx).$$

(c) Suppose in addition that $f = 0$ on the lower half-space $\{x_m < 0\}$. Show that $w := u - v$ solves

$$\begin{cases} \Delta w = f & \text{in } \mathbb{R}^m_+, \\ w = 0 & \text{on } \partial\mathbb{R}^m_+. \end{cases}$$

Solution. **(a)** Fix $x \in \mathbb{R}^m$. Near the singularity, for sufficiently small $\delta > 0$,

$$\begin{aligned}\int_{B(x,\delta)} \left|f(y)N(x-y)\right| \mathrm{d}y &\lesssim \left\|f\right\|_{L^\infty(B(x,\delta))} \int_{B(x,\delta)} \frac{\mathrm{d}y}{\left|x-y\right|^{m-2}} \\ &= C\int_0^\delta r^{2-m} r^{m-1}\,\mathrm{d}r \lesssim \delta^2.\end{aligned}$$

For $R > 2|x|$, the tail satisfies

$$\int_{\mathbb{R}^m \setminus B(0,R)} \left|f(y)N(x-y)\right| \mathrm{d}y \lesssim \int_R^\infty r^{-A} r^{2-m} r^{m-1}\,\mathrm{d}r$$
$$= \int_R^\infty r^{1-A}\,\mathrm{d}r = \frac{R^{2-A}}{A-2} < \infty.$$

The remaining region is compact and avoids the singularity, so $u(x)$ is well defined.

For the first derivatives,

$$\partial_i N(x) = c_m(2-m)\frac{x_i}{|x|^m}, \qquad |\partial_i N(x)| \lesssim \frac{1}{|x|^{m-1}}.$$

Consequently,

$$\int_{B(x,\delta)} \left|f(y)\partial_i N(x-y)\right| \mathrm{d}y \lesssim C\int_0^\delta \mathrm{d}r \lesssim \delta,$$

whereas at infinity

$$\int_R^\infty r^{-A} r^{1-m} r^{m-1}\,\mathrm{d}r = \int_R^\infty r^{-A}\,\mathrm{d}r = \frac{R^{1-A}}{A-1} < \infty.$$

Thus differentiation under the integral sign gives $u \in C^1$. Although

$$|\partial_{ii} N(x)| \lesssim |x|^{-m}$$

is not locally integrable, the Newtonian-potential theorem gives $\Delta u = f$; local elliptic regularity then yields $u \in C^\infty$ where f is smooth.

The displayed weighted estimate gives

$$|x|^B\,|u(x)| \lesssim |x|^{B+2-A} + |x|^{B+1}.$$

Its right-hand side tends to zero as $|x| \to \infty$ provided

$$B < A-2, \qquad B < -1,$$

and it tends to zero as $|x| \to 0$ provided the opposite inequalities hold:

$$B > A-2, \qquad B > -1.$$

(b) Since T is orthogonal, $|\det T| = 1$, and N is radial. Therefore, with $z = Ty$,

$$\begin{aligned} v(x) &= \int_{\mathbb{R}^m} f\big(T(x-y)\big)N(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^m} f(Tx-z)N(Tz)\,\mathrm{d}z \\ &= \int_{\mathbb{R}^m} f(Tx-z)N(z)\,\mathrm{d}z = u(Tx). \end{aligned}$$

(c) If $x_m = 0$, then $Tx = x$, and part (b) gives

$$w(x) = u(x) - v(x) = u(x) - u(Tx) = 0.$$

Moreover,

$$\Delta u = f, \qquad \Delta v = \widetilde{f} = f \circ T.$$

For $x_m > 0$, the point Tx lies in the lower half-space, where f vanishes. Hence $\widetilde{f}(x) = 0$, and

$$\Delta w = f - \widetilde{f} = f \qquad \text{in } \mathbb{R}^m_+.$$

□

Problem 4.3. Write $z = x + \mathrm{i}y$, and let

$$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}, \qquad \mathbb{D}^* := \mathbb{D} \setminus \{0\}.$$

(a) If f and g are analytic, show that

$$u(x, y) = f(z) + g(\overline{z})$$

is harmonic.

(b) Suppose

$$F(\theta) = \sum_{m=-\infty}^{\infty} c_m \mathrm{e}^{\mathrm{i}m\theta}, \qquad \sum_{m=-\infty}^{\infty} |c_m| < \infty.$$

Construct the harmonic solution of the Dirichlet problem in $\mathbb{D}$ with boundary value F, and prove uniqueness.

(c) Show that the same boundary-value problem is not unique in the punctured disk if no condition is imposed at the origin.

(d) Suppose w is harmonic in $\mathbb{D}^*$, vanishes on $\partial\mathbb{D}$, and satisfies, uniformly in θ,

$$\frac{w(r, \theta)}{\log(r^2)} \longrightarrow 0 \qquad \text{as } r \longrightarrow 0.$$

Prove that $w \equiv 0$.

(e) Prove that the solution constructed in part (b) is the unique solution on $\mathbb{D}^*$ satisfying the condition in part (d).

Solution. **(a)** Analyticity gives $f, g \in C^2$. Since

$$\partial_{xx} u = f''(z) + g''(\overline{z})$$

and

$$\partial_{yy} u = \mathrm{i}^2 f''(z) + (-\mathrm{i})^2 g''(\overline{z}) = -f''(z) - g''(\overline{z}),$$

we obtain $\Delta u = 0$.

(b) Define

$$u(r, \theta) := \sum_{m=-\infty}^{\infty} c_m r^{|m|} \mathrm{e}^{\mathrm{i}m\theta}.$$

Splitting the nonnegative and negative modes gives

$$u(r, \theta) = \sum_{m=0}^{\infty} c_m r^m \mathrm{e}^{\mathrm{i}m\theta} + \sum_{m=1}^{\infty} c_{-m} r^m \mathrm{e}^{-\mathrm{i}m\theta}$$

$$
\begin{aligned}
&= \sum_{m=0}^{\infty} c_m z^m + \sum_{m=1}^{\infty} c_{-m} \overline{z}^m \\
&= f(z) + g(\overline{z}).
\end{aligned}
$$

The absolute summability of the coefficients makes f and g analytic in $\mathbb{D}$, so part (a) shows that u is harmonic. Moreover,

$$u(1, \theta) = F(\theta).$$

The maximum principle applied to the difference of two solutions gives uniqueness in $C(\overline{\mathbb{D}}) \cap C^2(\mathbb{D})$.

(c) For any $\varepsilon \in \mathbb{R}$, set

$$u_\varepsilon(r, \theta) := u(r, \theta) + \varepsilon \log(r^2).$$

Then

$$u_\varepsilon(1, \theta) = F(\theta),$$

and

$$\partial_r u_\varepsilon = \partial_r u + \frac{2\varepsilon}{r}, \qquad \partial_{rr} u_\varepsilon = \partial_{rr} u - \frac{2\varepsilon}{r^2}.$$

Thus

$$\partial_{rr} u_\varepsilon + \frac{1}{r}\partial_r u_\varepsilon + \frac{1}{r^2}\partial_{\theta\theta} u_\varepsilon = 0 \qquad \text{in } \mathbb{D}^*.$$

Different values of ε therefore give different harmonic solutions with the same boundary value.

(d) Fix $\varepsilon > 0$. The hypothesis gives $\delta_\varepsilon \in (0, 1)$ such that

$$|w(\delta_\varepsilon, \theta)| \le \varepsilon \left|\log(\delta_\varepsilon^2)\right| \qquad \text{for every } \theta.$$

On the annulus $\delta_\varepsilon < r < 1$, the functions $\pm w(r, \theta) - \varepsilon \left|\log(r^2)\right|$ are harmonic. They are nonpositive on both boundary circles, so the maximum principle gives

$$|w(r, \theta)| \le \varepsilon \left|\log(r^2)\right| \qquad (\delta_\varepsilon \le r \le 1).$$

For fixed $r > 0$, choose each $\delta_\varepsilon < r$, as the limiting hypothesis permits, and let $\varepsilon \to 0$. Then $w(r, \theta) = 0$. Hence $w \equiv 0$ in $\mathbb{D}^*$.

(e) The solution in part (b) satisfies

$$|u(r, \theta)| \le \sum_{m=-\infty}^{\infty} |c_m| < \infty.$$

Therefore

$$\lim_{r \to 0} \frac{u(r, \theta)}{\log(r^2)} = 0$$

uniformly in θ. If v is another solution with the same boundary value and the same behavior at the origin, then $w = u - v$ satisfies part (d), so $u = v$. $\square$

Chapter 5

Sobolev Spaces, Fourier Analysis, and Elliptic Regularity

5.1 Sobolev spaces

Let $1 \le p \le \infty$ and let $k \in \mathbb{Z}_{\ge 0}$. The Sobolev space $W^{k,p}(\mathbb{R}^m)$ consists of all functions $u \in L^1_{\mathrm{loc}}(\mathbb{R}^m)$ such that the weak derivative $\partial^\alpha u$ exists and belongs to $L^p(\mathbb{R}^m)$ for every multi-index α with $|\alpha| \le k$.

The norm is

$$\|u\|_{W^{k,p}} := \begin{cases} \left(\sum_{|\alpha| \le k} \int_{\mathbb{R}^m} |\partial^\alpha u(x)|^p \, \mathrm{d}x \right)^{1/p}, & 1 \le p < \infty, \\ \sum_{|\alpha| \le k} \|\partial^\alpha u\|_{L^\infty}, & p = \infty. \end{cases} \tag{5.1}$$

Thus

$$u_n \longrightarrow u \quad \text{in } W^{k,p} \qquad \Longleftrightarrow \qquad \|u_n - u\|_{W^{k,p}} \longrightarrow 0.$$

Theorem 5.1 (Completeness of Sobolev spaces)**.** *For every $k \in \mathbb{Z}_{\ge 0}$ and $1 \le p \le \infty$, $W^{k,p}(\mathbb{R}^m)$ is a Banach space.*

Proof. The expression in (5.1) is a norm. Let (u_n) be Cauchy in $W^{k,p}$. For every $|\alpha| \le k$, the sequence $(\partial^\alpha u_n)$ is Cauchy in L^p. Completeness of L^p gives a function $u_\alpha \in L^p$ such that

$$\partial^\alpha u_n \longrightarrow u_\alpha \quad \text{in } L^p.$$

In particular, for the zero multi-index,

$$u_n \longrightarrow u_0 =: u \quad \text{in } L^p.$$

It remains to identify u_α as the weak derivative of u. Take $\varphi \in C_c^\infty(\mathbb{R}^m)$. Strong L^p convergence implies convergence against this test function, and therefore

$$\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle$$

$$
\begin{aligned}
&= (-1)^{|\alpha|} \lim_{n\to\infty} \langle u_n, \partial^\alpha \varphi\rangle \\
&= \lim_{n\to\infty} \langle \partial^\alpha u_n, \varphi\rangle = \langle u_\alpha, \varphi\rangle.
\end{aligned}
$$

Hence $\partial^\alpha u = u_\alpha \in L^p$ for every $|\alpha| \le k$, so $u \in W^{k,p}$. The convergence of all these derivatives gives $u_n \to u$ in $W^{k,p}$. □

5.1.1 Approximation by smooth functions

Theorem 5.2 (Density of compactly supported smooth functions)**.** *If $1 \le p < \infty$, then $C_c^\infty(\mathbb{R}^m)$ is dense in $W^{k,p}(\mathbb{R}^m)$.*

Proof. Choose a standard mollifier $\varphi \in C_c^\infty(\mathbb{R}^m)$ and set

$$
\varphi_\varepsilon(x) = \varepsilon^{-m}\varphi(x/\varepsilon), \qquad u_\varepsilon = u * \varphi_\varepsilon.
$$

Then $u_\varepsilon \in C^\infty(\mathbb{R}^m)$. For $|\alpha| \le k$, differentiation of the convolution, followed by the definition of a weak derivative, gives

$$
\begin{aligned}
\partial^\alpha u_\varepsilon(x) &= \int_{\mathbb{R}^m} u(y)\,\partial_x^\alpha \varphi_\varepsilon(x-y)\,\mathrm{d}y \\
&= (-1)^{|\alpha|}\int_{\mathbb{R}^m} u(y)\,\partial_y^\alpha \varphi_\varepsilon(x-y)\,\mathrm{d}y \\
&= \int_{\mathbb{R}^m} \partial^\alpha u(y)\,\varphi_\varepsilon(x-y)\,\mathrm{d}y = (\partial^\alpha u) * \varphi_\varepsilon(x).
\end{aligned}
$$

Since $\partial^\alpha u \in L^p$,

$$
\partial^\alpha u_\varepsilon \longrightarrow \partial^\alpha u \quad \text{in } L^p \qquad (|\alpha| \le k).
$$

Consequently $u_\varepsilon \to u$ in $W^{k,p}$.

Mollification makes a function smooth, but does not give compact support when u is not compactly supported. We therefore add a cutoff. Choose $\rho \in C_c^\infty(\mathbb{R}^m)$ with $\rho = 1$ near the origin, and set

$$
\rho_R(x) = \rho(x/R), \qquad R \longrightarrow \infty.
$$

If $v \in C^\infty(\mathbb{R}^m) \cap W^{k,p}(\mathbb{R}^m)$, then $v_R = v\rho_R \in C_c^\infty(\mathbb{R}^m)$. For the first derivatives,

$$
\partial_j v_R = (\partial_j v)\rho_R + v\,\partial_j \rho_R.
$$

Dominated convergence gives $(\partial_j v)\rho_R \to \partial_j v$ in L^p, while $v\,\partial_j\rho_R \to 0$ in L^p. Applying the same argument to the higher-order Leibniz formula yields $v_R \to v$ in $W^{k,p}$. Combining this cutoff with the mollification completes the proof. □

Corollary 5.3. *The Schwartz space $\mathcal{S}(\mathbb{R}^m)$ is dense in $W^{k,p}(\mathbb{R}^m)$ for $1 \le p < \infty$.*

Proof. This follows immediately from

$$
C_c^\infty(\mathbb{R}^m) \subset \mathcal{S}(\mathbb{R}^m) \subset W^{k,p}(\mathbb{R}^m).
$$

□

Remark 5.4. The notation

$$H^1(\mathbb{R}^m) = W^{1,2}(\mathbb{R}^m), \qquad H^k(\mathbb{R}^m) = W^{k,2}(\mathbb{R}^m)$$

is standard. These are Hilbert spaces. For instance,

$$\langle u, v\rangle_{H^1} = \int_{\mathbb{R}^m} \left(u\overline{v} + \nabla u \cdot \overline{\nabla v} \right) \, \mathrm{d}x.$$

5.2 The critical Sobolev exponent

Assume throughout this section that $1 \le p < m$. We seek an estimate of the form

$$\|u\|_{L^q} \le C\|\nabla u\|_{L^p}, \qquad u \in C_c^\infty(\mathbb{R}^m), \tag{5.2}$$

where C does not depend on u.

5.2.1 Scaling motivation

If an inequality such as (5.2) is possible, scaling forces a specific value of q. Fix a nonzero $u \in C_c^\infty(\mathbb{R}^m)$ and define

$$u_\lambda(x) = u(\lambda x), \qquad \lambda > 0.$$

Changing variables gives

$$\int_{\mathbb{R}^m} |u_\lambda|^q \, \mathrm{d}x = \lambda^{-m} \int_{\mathbb{R}^m} |u|^q \, \mathrm{d}y,$$

and

$$\begin{aligned} \int_{\mathbb{R}^m} |\nabla u_\lambda|^p \, \mathrm{d}x &= \lambda^p \int_{\mathbb{R}^m} |\nabla u(\lambda x)|^p \, \mathrm{d}x \\ &= \lambda^{p-m} \int_{\mathbb{R}^m} |\nabla u(y)|^p \, \mathrm{d}y. \end{aligned}$$

Thus (5.2) applied to u_λ would imply

$$\|u\|_{L^q} \le C\lambda^{1-m/p+m/q}\|\nabla u\|_{L^p}.$$

If $1 - m/p + m/q \ne 0$, sending λ to either 0 or ∞ gives a contradiction. Therefore

$$\frac{1}{q} = \frac{1}{p} - \frac{1}{m}.$$

Definition 5.5. For $1 \le p < m$, the *Sobolev conjugate* of p is

$$p^* = \frac{mp}{m-p}.$$

Equivalently,

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{m}, \qquad p^* > p.$$

The scaling calculation shows that (5.2) can hold only for $q = p^*$.

5.2.2 The Gagliardo–Nirenberg–Sobolev inequality

Theorem 5.6 (Gagliardo–Nirenberg–Sobolev inequality)**.** *Let $1 \le p < m$. There is a constant $C = C(m,p)$ such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^m)} \le C\|\nabla u\|_{L^p(\mathbb{R}^m)} \tag{5.3}$$

for every $u \in C_c^\infty(\mathbb{R}^m)$.

Remark 5.7. Compact support is essential in this homogeneous form. The constant function $u \equiv 1$ has zero gradient but does not belong to $L^{p^*}(\mathbb{R}^m)$.

Proof. We begin with $p = 1$ and illustrate the argument in dimension $m = 3$. Since u has compact support, the fundamental theorem of calculus gives

$$\begin{aligned} |u(x_1, x_2, x_3)| &\le \int_{-\infty}^{x_1} |\partial_1 u(s_1, x_2, x_3)| \, \mathrm{d}s_1 \\ &\le \int_{\mathbb{R}} |\nabla u(s_1, x_2, x_3)| \, \mathrm{d}s_1. \end{aligned}$$

The corresponding estimates in the x_2 and x_3 directions also hold. Define

$$\begin{aligned} A_1(x_2, x_3) &= \int_{\mathbb{R}} |\nabla u(s_1, x_2, x_3)| \, \mathrm{d}s_1, \\ A_2(x_1, x_3) &= \int_{\mathbb{R}} |\nabla u(x_1, s_2, x_3)| \, \mathrm{d}s_2, \\ A_3(x_1, x_2) &= \int_{\mathbb{R}} |\nabla u(x_1, x_2, s_3)| \, \mathrm{d}s_3. \end{aligned}$$

Multiplication of the three one-dimensional estimates yields

$$|u(x_1, x_2, x_3)|^{3/2} \le A_1(x_2, x_3)^{1/2} A_2(x_1, x_3)^{1/2} A_3(x_1, x_2)^{1/2}.$$

Integrate first in x_1, then in x_2, and finally in x_3. Repeated use of Hölder's inequality with exponent 2, together with Fubini's theorem, gives

$$\int_{\mathbb{R}^3} |u|^{3/2} \, \mathrm{d}x \le \left(\int_{\mathbb{R}^3} |\nabla u| \, \mathrm{d}x \right)^{3/2}.$$

Therefore

$$\|u\|_{L^{3/2}(\mathbb{R}^3)} \le \|\nabla u\|_{L^1(\mathbb{R}^3)}.$$

The same argument in $\mathbb{R}^m$, using generalized Hölder inequality with exponent $m-1$, gives

$$\|u\|_{L^{m/(m-1)}} \le C_m \|\nabla u\|_{L^1}. \tag{5.4}$$

Now let $1 < p < m$ and apply (5.4) to $|u|^\gamma$, with $\gamma > 1$. We obtain

$$\left(\int_{\mathbb{R}^m} |u|^{\gamma m/(m-1)} \, \mathrm{d}x \right)^{(m-1)/m} \le C_m \int_{\mathbb{R}^m} |\nabla(|u|^\gamma)| \, \mathrm{d}x$$

$$\begin{aligned}&\le C_m \gamma \int_{\mathbb{R}^m} |u|^{\gamma-1} |\nabla u| \, \mathrm{d}x \\ &\le C_m \gamma \left(\int_{\mathbb{R}^m} |u|^{(\gamma-1)p/(p-1)} \, \mathrm{d}x \right)^{(p-1)/p} \|\nabla u\|_{L^p}.\end{aligned}$$

Choose γ so that the two powers of $|u|$ agree:

$$\frac{(\gamma-1)p}{p-1} = \frac{\gamma m}{m-1}.$$

This gives

$$\gamma = \frac{p(m-1)}{m-p} > 1, \qquad \frac{\gamma m}{m-1} = \frac{mp}{m-p} = p^*.$$

After cancelling the resulting power of $\|u\|_{L^{p^*}}$, we find

$$\|u\|_{L^{p^*}} \le C_m \frac{p(m-1)}{m-p} \|\nabla u\|_{L^p},$$

which is (5.3) with a constant depending only on m and p. □

5.2.3 Interpolation consequences

Corollary 5.8. *Let $1 \le p < m$, $1 \le q \le \infty$, and $0 \le \theta \le 1$. If*

$$\frac{1}{r} = \frac{\theta}{p^*} + \frac{1-\theta}{q},$$

then

$$\|u\|_{L^r} \le C^\theta \|\nabla u\|_{L^p}^\theta \|u\|_{L^q}^{1-\theta} \tag{5.5}$$

for every $u \in C_c^\infty(\mathbb{R}^m)$, with the usual interpretation at the endpoints.

Proof. Since $1 = r\theta/p^* + r(1-\theta)/q$, Hölder's inequality gives

$$\begin{aligned}\int_{\mathbb{R}^m} |u|^r \, \mathrm{d}x &= \int_{\mathbb{R}^m} |u|^{r\theta} |u|^{r(1-\theta)} \, \mathrm{d}x \\ &\le \|u\|_{L^{p^*}}^{r\theta} \|u\|_{L^q}^{r(1-\theta)}.\end{aligned}$$

Take the rth root and apply (5.3). □

The notes also record a higher-order Gagliardo–Nirenberg interpolation inequality, although it will not be used here. Let the ambient dimension be n, let $0 \le j < \ell$, and write

$$\partial^{(j)} u = (\partial^\alpha u)_{|\alpha|=j}, \qquad \partial^{(\ell)} u = (\partial^\alpha u)_{|\alpha|=\ell}.$$

If

$$\frac{1}{r} = (1-\theta)\left(\frac{1}{q} - \frac{j}{n}\right) + \theta\left(\frac{1}{p} - \frac{\ell}{n}\right), \qquad \frac{j}{\ell} \le \theta < 1,$$

then, under the standard admissibility conditions on the exponents,

$$\|u\|_{L^r} \le C\|\partial^{(j)}u\|_{L^q}^{1-\theta}\|\partial^{(\ell)}u\|_{L^p}^{\theta}. \tag{5.6}$$

In particular, the first Sobolev inequality gives the continuous embedding

$$W^{1,p}(\mathbb{R}^m) \hookrightarrow L^{p^*}(\mathbb{R}^m), \qquad 1 \le p < m,$$

after approximation by compactly supported smooth functions.

Example 5.9 (The case $m = 3$ and $p = 2$)**.** Here $p^* = 6$, so

$$\|u\|_{L^6(\mathbb{R}^3)} \lesssim \|\nabla u\|_{L^2(\mathbb{R}^3)}.$$

Suppose $u_n \in C_c^\infty(\mathbb{R}^3)$ and $\sup_n \|\nabla u_n\|_{L^2} < \infty$. If u_n converged almost everywhere to a nonzero constant c, Fatou's lemma would give

$$\int_{\mathbb{R}^3} |c|^6 \, \mathrm{d}x \le \liminf_{n\to\infty} \|u_n\|_{L^6}^6 \lesssim \liminf_{n\to\infty} \|\nabla u_n\|_{L^2}^6 < \infty,$$

a contradiction. Hence such a sequence cannot converge almost everywhere to a nonzero constant.

If in addition $\|\nabla u_n\|_{L^2} \to 0$, then $u_n \to 0$ in L^6. Consequently a subsequence converges to zero almost everywhere.

5.3 Schwartz functions and tempered distributions

Definition 5.10. The Schwartz space is

$$\mathcal{S}(\mathbb{R}^m) = \left\{\varphi \in C^\infty(\mathbb{R}^m) : N_r(\varphi) < \infty \text{ for every } r \ge 0\right\},$$

where

$$N_r(\varphi) := \sup_{x\in\mathbb{R}^m} \sum_{|\alpha|+|\beta|\le r} |x^\alpha \partial^\beta \varphi(x)|.$$

Thus a Schwartz function is smooth and, together with all its derivatives, decays faster than any polynomial as $|x| \to \infty$. Since

$$C_c^\infty(\mathbb{R}^m) \subset \mathcal{S}(\mathbb{R}^m),$$

duality reverses the inclusion:

$$\mathcal{S}'(\mathbb{R}^m) \subset \mathcal{D}'(\mathbb{R}^m).$$

The elements of $\mathcal{S}'$ are called *tempered distributions*.

A distribution $T \in \mathcal{D}'$ has, on every ball, an estimate of the form

$$|\langle T, \varphi\rangle| \le C_R\|\varphi\|_{C^N}, \qquad \varphi \in C_c^\infty(B_R(0)).$$

If the order N can be chosen independently of R and C_R grows at most polynomially in R, then T extends continuously to $\mathcal{S}$ and therefore belongs to $\mathcal{S}'$.

Example 5.11. Every $f \in L^1_{\mathrm{loc}}(\mathbb{R}^m)$ defines a distribution in $\mathcal{D}'$. If f grows at most polynomially, it defines a tempered distribution. Exponential growth need not be tempered: for example, e^x on $\mathbb{R}$ belongs to $\mathcal{D}'$ but not to $\mathcal{S}'$.

5.3.1 The Fourier transform on the Schwartz space

For $f \in \mathcal{S}(\mathbb{R}^m)$, define

$$\widehat{f}(\xi) = \int_{\mathbb{R}^m} f(x) e^{-2\pi \mathrm{i} x\cdot\xi}\, \mathrm{d}x. \tag{5.7}$$

Integration by parts and differentiation under the integral sign give

$$\widehat{\partial_k f}(\xi) = 2\pi\mathrm{i}\,\xi_k \widehat{f}(\xi), \tag{5.8}$$

$$\widehat{x_j f}(\xi) = -\frac{1}{2\pi\mathrm{i}}\,\partial_{\xi_j}\widehat{f}(\xi). \tag{5.9}$$

More generally,

$$\widehat{x^\alpha \partial^\beta f}(\xi) = \left(-\frac{1}{2\pi\mathrm{i}}\right)^{|\alpha|} \partial_\xi^\alpha \left((2\pi\mathrm{i}\xi)^\beta \widehat{f}(\xi)\right). \tag{5.10}$$

These identities show that

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^m) \longrightarrow \mathcal{S}(\mathbb{R}^m).$$

The inverse transform is

$$f^\vee(x) = \int_{\mathbb{R}^m} e^{2\pi\mathrm{i} x\cdot\xi} f(\xi)\, \mathrm{d}\xi, \tag{5.11}$$

and Fourier inversion states

$$(\widehat{f})^\vee = f.$$

With respect to the L^2 inner product,

$$\mathcal{F}^* = \mathcal{F}^{-1}.$$

For $a > 0$, the Gaussian is transformed into another Gaussian:

$$\widehat{e^{-\pi a|x|^2}}(\xi) = a^{-m/2} e^{-\pi|\xi|^2/a}. \tag{5.12}$$

Plancherel's theorem gives

$$\|f\|_{L^2} = \|\widehat{f}\|_{L^2}. \tag{5.13}$$

The Fourier transform consequently has the following extensions:

$$\mathcal{F} : L^1 \longrightarrow L^\infty, \qquad \mathcal{F} : L^2 \longrightarrow L^2, \qquad \mathcal{F} : \mathcal{S}' \longrightarrow \mathcal{S}'.$$

The second extension follows from Plancherel and density of $\mathcal{S}$ in L^2; the third is defined by duality.

5.4 Fourier-defined Sobolev spaces

For $u \in H^1(\mathbb{R}^m)$, Plancherel and (5.8) give

$$\int_{\mathbb{R}^m} \left(1 + 4\pi^2|\xi|^2\right) |\widehat{u}(\xi)|^2\, \mathrm{d}\xi < \infty. \tag{5.14}$$

The factor $4\pi^2$ is harmless for norm equivalence. This suggests the following definition for every real order.

Definition 5.12. For $s \in \mathbb{R}$, the space $H^s(\mathbb{R}^m)$ consists of those $u \in \mathcal{S}'(\mathbb{R}^m)$ for which $\widehat{u}$ is represented by a locally square-integrable function and

$$\|u\|_{H^s}^2 := \int_{\mathbb{R}^m} (1+|\xi|^2)^s |\widehat{u}(\xi)|^2 \, \mathrm{d}\xi < \infty. \tag{5.15}$$

When $s < 0$, elements of H^s need not be functions; the definition is therefore made inside the tempered distributions. When $s \geq 0$, the weight is at least one, so every element of H^s is represented by an L^2 function. If $s = k \in \mathbb{Z}_{\geq 0}$, then

$$H^k(\mathbb{R}^m) = W^{k,2}(\mathbb{R}^m)$$

with equivalent norms. Each H^s is a Hilbert space: changing s only changes the Fourier-side measure by the positive weight $(1+|\xi|^2)^s$.

The dual of H^s is H^{-s}. Indeed, if $u \in H^{-s}$ and $\varphi \in H^s$, then the Fourier-side pairing and Cauchy–Schwarz give

$$\begin{aligned} |\langle u, \varphi\rangle| &\leq \int_{\mathbb{R}^m} |\widehat{u}(\xi)|, |\widehat{\varphi}(\xi)| \, \mathrm{d}\xi \\ &\leq \left(\int_{\mathbb{R}^m} (1+|\xi|^2)^{-s} |\widehat{u}|^2 \, \mathrm{d}\xi\right)^{1/2} \left(\int_{\mathbb{R}^m} (1+|\xi|^2)^{s} |\widehat{\varphi}|^2 \, \mathrm{d}\xi\right)^{1/2} \\ &= \|u\|_{H^{-s}} \|\varphi\|_{H^s}. \end{aligned}$$

Thus

$$(H^s)' = H^{-s}.$$

Remark 5.13 (Polynomial growth). If P is a polynomial of degree ℓ, then

$$|P(\xi)| \leq C_1 + C_2 |\xi|^\ell.$$

A matching lower estimate $c_1|\xi|^\ell - c_2 \leq |P(\xi)|$ is not valid for every polynomial in several variables. It holds when the leading homogeneous part is nonzero in every direction, which is the elliptic case used in Fourier estimates for constant-coefficient differential operators.

5.5 Applications to elliptic equations

5.5.1 A Fourier-transform proof of Liouville's theorem

Classically, a bounded harmonic function on all of $\mathbb{R}^m$ is constant. Fourier analysis gives another proof for a tempered distributional solution. Suppose

$$\Delta u = 0 \quad \text{in } \mathcal{S}'(\mathbb{R}^m).$$

Then

$$\widehat{\partial_i^2 u}(\xi) = -(2\pi)^2 \xi_i^2 \widehat{u}(\xi),$$

and hence

$$\widehat{\Delta u}(\xi) = -(2\pi)^2 |\xi|^2 \widehat{u}(\xi) = 0.$$

It follows that $\widehat{u}$ is supported at the origin. A distribution supported at one point is a finite sum of derivatives of the Dirac mass:

$$\widehat{u} = \sum_{|\alpha|\le N} c_\alpha \partial^\alpha \delta_0.$$

Taking the inverse Fourier transform shows that u is a polynomial. The equation says that this polynomial is harmonic, and a bounded polynomial is constant. This recovers Liouville's theorem.

5.5.2 A global H^2 estimate

Consider

$$(1-\Delta)u = f, \qquad f \in L^2(\mathbb{R}^m), \quad u \in \mathcal{S}'(\mathbb{R}^m). \tag{5.16}$$

Taking Fourier transforms gives

$$\left(1 + 4\pi^2|\xi|^2\right)\widehat{u}(\xi) = \widehat{f}(\xi),$$

and therefore

$$\widehat{u}(\xi) = \frac{\widehat{f}(\xi)}{1+4\pi^2|\xi|^2}. \tag{5.17}$$

Since $1+|\xi|^2$ is comparable to $1+4\pi^2|\xi|^2$, we obtain

$$\begin{aligned}\|u\|_{H^2}^2 &= \int_{\mathbb{R}^m} (1+|\xi|^2)^2 |\widehat{u}(\xi)|^2 \, \mathrm{d}\xi \\ &\le C \int_{\mathbb{R}^m} |\widehat{f}(\xi)|^2 \, \mathrm{d}\xi = C\|f\|_{L^2}^2.\end{aligned}$$

Thus $u \in H^2(\mathbb{R}^m)$ and

$$\|u\|_{H^2} \le C\|f\|_{L^2}. \tag{5.18}$$

5.5.3 Localization

On a bounded domain, the global Fourier transform is not directly available. A cutoff reduces an interior question to the global estimate above. Suppose

$$\Delta u = f \quad \text{in } \Omega, \qquad f \in L^2(\Omega), \quad u \in H^1(\Omega).$$

We want to prove that $u \in H^2_{\mathrm{loc}}(\Omega)$.

Choose $x_0 \in \Omega$ and a smooth cutoff $\varphi \in C_c^\infty(\Omega)$ which equals 1 near x_0.

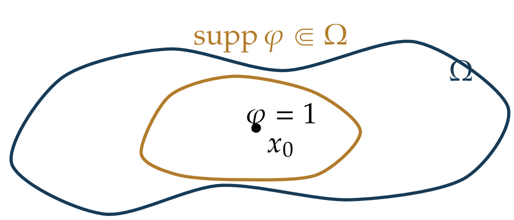


Figure 5.1: A cutoff equal to one near x_0 localizes the equation inside Ω.

The product rule gives

$$\Delta(u\varphi) = f\varphi + 2\nabla u \cdot \nabla\varphi + u\Delta\varphi. \tag{5.19}$$

Equivalently,

$$(\Delta - 1)(u\varphi) = 2\nabla u \cdot \nabla\varphi + u\Delta\varphi + f\varphi - u\varphi. \tag{5.20}$$

Every term on the right belongs to $L^2(\mathbb{R}^m)$ after extension by zero, because $u \in H^1$, $f \in L^2$, and φ is smooth and compactly supported in Ω. Applying the global H^2 estimate to $u\varphi$ gives

$$u\varphi \in H^2(\mathbb{R}^m).$$

Since $\varphi = 1$ near x_0, we conclude that

$$u \in H^2_{\mathrm{loc}}(\Omega).$$

5.6 Problems

Problem 5.1 (Order of some distributions)**.** Determine the order of the following distributions.

(a) The Dirac mass δ_0 on $\mathbb{R}$.
(b) The principal-value distribution $\mathrm{p.\,v.}(1/x)$ on $\mathbb{R}$.
(c) Let $(x_m)_{m\geq 1}$ be a sequence of distinct points in $(0,1)$ such that $x_m \longrightarrow 0$, and define

$$T = \sum_{m=1}^{\infty} \partial_x^m \delta_{x_m}.$$

Show that T is a distribution on $(0,1)$ but has no finite order.

Solution. For (a), if $\phi \in C_c^\infty(\mathbb{R})$, then

$$\left|\langle \delta_0, \phi\rangle\right| = \left|\phi(0)\right| \leq \left\|\phi\right\|_{L^\infty}.$$

Thus δ_0 has order zero.

For (b),

$$\left\langle \mathrm{p.\,v.}(1/x), \phi\right\rangle = \int_0^\infty \frac{\phi(x) - \phi(-x)}{x}\,\mathrm{d}x.$$

If $\operatorname{supp}\phi \subset [-\ell, \ell]$, the mean-value theorem gives

$$\left|\left\langle \mathrm{p.\,v.}(1/x), \phi\right\rangle\right| \leq 2\ell\left\|\partial_x\phi\right\|_{L^\infty}.$$

Hence $\mathrm{p.\,v.}(1/x)$ has order at most one. To see that its order is not zero, choose $\phi_\varepsilon \in C_c^\infty((0,2))$ such that

$$0 \leq \phi_\varepsilon \leq 1, \qquad \phi_\varepsilon = 1 \quad \text{on } [2\varepsilon, 1].$$

Then $\left\|\phi_\varepsilon\right\|_{L^\infty} = 1$, whereas

$$\left\langle \mathrm{p.\,v.}(1/x), \phi_\varepsilon\right\rangle \geq \int_{2\varepsilon}^1 \frac{\mathrm{d}x}{x} = \log\frac{1}{2\varepsilon} \longrightarrow \infty.$$

Therefore p. v.$(1/x)$ has order exactly one.

For (c), every compact set $K \Subset (0,1)$ contains only finitely many of the points x_m. Consequently, for $\phi \in C_c^\infty(K)$,

$$\langle T, \phi \rangle = \sum_{x_m \in K} (-1)^m \partial_x^m \phi(x_m),$$

and this finite sum is bounded by a constant times a finite C^N seminorm of ϕ. Thus $T \in \mathcal{D}'((0,1))$.

The distribution $\partial_x^m \delta_{x_m}$ has order exactly m. It has order at most m by its definition. If it had order at most $m-1$, take $\eta \in C_c^\infty((-1,1))$ with $\partial_x^m \eta(0) \neq 0$ and set

$$\eta_\varepsilon(x) = \varepsilon^m \eta\left(\frac{x - x_m}{\varepsilon}\right).$$

For sufficiently small ε, these functions have support in a fixed compact neighborhood of x_m; their derivatives through order $m-1$ tend uniformly to zero, but $\partial_x^m \eta_\varepsilon(x_m) = \partial_x^m \eta(0)$. This is a contradiction. Finally, a small neighborhood of x_m containing no other x_j restricts T to $\partial_x^m \delta_{x_m}$. Since m is arbitrary, T has no finite order. ☐

Problem 5.2 (Distributional antiderivatives)**.** Let $I = (a,b)$ and let $f \in L^1(I)$.

(a) Show that if $T \in \mathcal{D}'(I)$ and $\partial_x T = 0$, then T is a constant distribution.
(b) Find all distributional solutions of $\partial_x T = f$ on I.

Solution. For (a), a distribution on an interval whose distributional derivative vanishes is constant.

For (b), set

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

For every $\phi \in C_c^\infty(I)$, Fubini's theorem and the fundamental theorem of calculus give

$$\begin{aligned}
\langle \partial_x F, \phi \rangle &= -\int_a^b \left(\int_a^x f(t)\,\mathrm{d}t \right) \partial_x \phi(x)\,\mathrm{d}x \\
&= -\int_a^b f(t) \left(\int_t^b \partial_x \phi(x)\,\mathrm{d}x \right) \mathrm{d}t \\
&= -\int_a^b f(t) \big(\phi(b) - \phi(t)\big)\,\mathrm{d}t \\
&= \int_a^b f(t)\phi(t)\,\mathrm{d}t = \langle f, \phi \rangle,
\end{aligned}$$

because $\phi(b) = 0$. Thus $\partial_x F = f$. If T is any other solution, then $\partial_x(T - F) = 0$, so part (a) yields

$$T = c + F, \qquad c \in \mathbb{R}.$$

These are all the solutions. ☐

Problem 5.3 (Difference quotients and mollification)**.** Let $1 < p < \infty$ and let $u \in W_0^{1,p}(0,1)$. Extend u by zero to $\mathbb{R}$, and for $h \neq 0$ define

$$D_h u(x) = \frac{u(x+h) - u(x)}{h}.$$

Let $\rho \in C_c^\infty((-1,1))$ be nonnegative with $\int_{\mathbb{R}} \rho = 1$, and put $\rho_\varepsilon(x) = \varepsilon^{-1}\rho(x/\varepsilon)$.

(a) Prove that, for $w \in W^{1,p}(\mathbb{R})$,

$$\|D_h w\|_{L^p(\mathbb{R})} \leq \|\partial_x w\|_{L^p(\mathbb{R})}.$$

(b) Prove that $u * \rho_\varepsilon \to u$ in $W^{1,p}(\mathbb{R})$ as $\varepsilon \to 0$.
(c) Prove that $D_h u \to \partial_x u$ in $L^p(\mathbb{R})$ as $h \to 0$.

Solution. For (a), first take w smooth. The fundamental theorem of calculus gives

$$D_h w(x) = \int_0^1 \partial_x w(x + th)\, \mathrm{d}t.$$

Equivalently, pairing with a test function and then using density in the dual space gives

$$\left|\left\langle D_h w, g\right\rangle\right| \leq \|\partial_x w\|_{L^p} \left\|g\right\|_{L^{p'}}, \qquad \frac{1}{p} + \frac{1}{p'} = 1.$$

Therefore, by duality,

$$\|D_h w\|_{L^p} \leq \|\partial_x w\|_{L^p}.$$

Approximation extends the estimate to every $w \in W^{1,p}(\mathbb{R})$.

For (b), a change of variables gives

$$(u * \rho_\varepsilon)(x) = \int_{\mathbb{R}} \rho(z) u(x - \varepsilon z)\, \mathrm{d}z,$$

and hence

$$\left|(u * \rho_\varepsilon)(x) - u(x)\right| \leq \int_{\mathbb{R}} \rho(z) \left|u(x - \varepsilon z) - u(x)\right|\, \mathrm{d}z.$$

By Minkowski's integral inequality,

$$\left\|u * \rho_\varepsilon - u\right\|_{L^p} \leq \int_{\mathbb{R}} \rho(z) \left\|u(cdot - \varepsilon z) - u\right\|_{L^p}\, \mathrm{d}z.$$

The integrand tends to zero by continuity of translations in L^p and is bounded by $2\rho(z)\|u\|_{L^p}$. Dominated convergence therefore yields

$$u * \rho_\varepsilon \longrightarrow u \quad \text{in } L^p(\mathbb{R}).$$

Moreover, differentiation commutes with convolution in the distributional sense:

$$\partial_x(u * \rho_\varepsilon) = u * (\partial_x \rho_\varepsilon) = (\partial_x u) * \rho_\varepsilon.$$

Applying the same L^p argument to $\partial_x u$ proves convergence in $W^{1,p}(\mathbb{R})$.

For (c), choose $u_m \in C_c^\infty((0,1))$ such that $u_m \to u$ in $W^{1,p}(\mathbb{R})$. For fixed m,

$$D_h u_m \longrightarrow \partial_x u_m \quad \text{in } L^p(\mathbb{R}) \qquad (h \to 0).$$

Using part (a),

$$\begin{aligned}
\|D_h u - \partial_x u\|_{L^p} &\leq \|D_h(u - u_m)\|_{L^p} + \|D_h u_m - \partial_x u_m\|_{L^p} + \|\partial_x u_m - \partial_x u\|_{L^p} \\
&\leq 2\,\|\partial_x u_m - \partial_x u\|_{L^p} + \|D_h u_m - \partial_x u_m\|_{L^p}.
\end{aligned}$$

First choose m large and then let $h \to 0$. It follows that

$$D_h u \longrightarrow \partial_x u \quad \text{in } L^p(\mathbb{R}).$$

$\square$

Problem 5.4 (Periodic Poincaré inequality). Let $L > 0$, and set

$$H^1_{\mathrm{per}}(0, L) := \left\{ u \in H^1(0, L) : u(0) = u(L) \right\}.$$

(a) If $u \in H^1_{\mathrm{per}}(0, L)$ and

$$\frac{1}{L} \int_0^L u(x)\,\mathrm{d}x = 0,$$

show that

$$\|u\|_{L^2(0,L)} \le \frac{L}{2\pi} \|\partial_x u\|_{L^2(0,L)}.$$

(b) For $u \in H^1_{\mathrm{per}}(0, L)$, define

$$u_{\mathrm{av}} := \frac{1}{L} \int_0^L u(x)\,\mathrm{d}x.$$

Show that

$$\|u - u_{\mathrm{av}}\|_{L^2(0,L)} \le \frac{L}{2\pi} \|\partial_x u\|_{L^2(0,L)}.$$

Solution. For part (a), first let $u \in C^\infty_{\mathrm{per}}([0, L])$ have zero average. With

$$e_m(x) := \mathrm{e}^{2\pi \mathrm{i} m x/L},$$

write

$$u(x) = \sum_{m \in \mathbb{Z}} a_m e_m(x), \qquad \partial_x u(x) = \sum_{m \in \mathbb{Z}} c_m e_m(x).$$

Both series converge uniformly. Their coefficients are

$$a_m = \frac{1}{L} \int_0^L u(x) \mathrm{e}^{-2\pi \mathrm{i} m x/L}\,\mathrm{d}x, \qquad c_m = \frac{1}{L} \int_0^L \partial_x u(x) \mathrm{e}^{-2\pi \mathrm{i} m x/L}\,\mathrm{d}x.$$

Integration by parts and periodicity give

$$\begin{aligned} c_m &= \frac{1}{L} \left[u(x) \mathrm{e}^{-2\pi \mathrm{i} m x/L} \right]_0^L + \frac{2\pi \mathrm{i} m}{L^2} \int_0^L u(x) \mathrm{e}^{-2\pi \mathrm{i} m x/L}\,\mathrm{d}x \\ &= \frac{2\pi \mathrm{i} m}{L} a_m. \end{aligned}$$

Since $a_0 = 0$, Plancherel's identity yields

$$\begin{aligned} \frac{1}{L} \int_0^L |u(x)|^2\,\mathrm{d}x &= \sum_{m \neq 0} |a_m|^2 \\ &\le \sum_{m \neq 0} m^2 |a_m|^2 \\ &= \frac{L^2}{4\pi^2} \sum_{m \in \mathbb{Z}} |c_m|^2 = \frac{L^2}{4\pi^2} \frac{1}{L} \int_0^L |\partial_x u(x)|^2\,\mathrm{d}x. \end{aligned}$$

The claimed estimate follows. Approximation by smooth periodic functions, with the averages subtracted from the approximants, extends the estimate to every zero-average $u \in H^1_{\mathrm{per}}(0, L)$.

For part (b), let

$$\widetilde{u} := u - u_{\mathrm{av}}.$$

Then $\widetilde{u}$ has zero average and $\partial_x \widetilde{u} = \partial_x u$. Part (a) therefore gives

$$\|u - u_{\mathrm{av}}\|_{L^2(0,L)} = \|\widetilde{u}\|_{L^2(0,L)} \le \frac{L}{2\pi} \|\partial_x u\|_{L^2(0,L)} .$$

□

Problem 5.5 (Uniform convergence of a sine series)**.** Let $(a_m)_{m\ge1} \subset \mathbb{C}$ satisfy

$$\sum_{m=1}^{\infty} m^2 |a_m|^2 < \infty,$$

and define

$$u_N(x) := \sum_{m=1}^{N} a_m \sin\left(\frac{2\pi m x}{L}\right), \qquad 0 \le x \le L.$$

Show that (u_N) converges uniformly on $[0, L]$ to a continuous function u, and that $u(0) = u(L) = 0$.

Solution. By the Cauchy–Schwarz inequality,

$$\sum_{m=1}^{\infty} |a_m| \le \left(\sum_{m=1}^{\infty} m^2 |a_m|^2\right)^{1/2} \left(\sum_{m=1}^{\infty} \frac{1}{m^2}\right)^{1/2} < \infty.$$

Consequently, if $N > M$, then

$$\sup_{x\in[0,L]} |u_N(x) - u_M(x)| \le \sum_{m=M+1}^{N} |a_m| .$$

The right-hand side tends to zero as $M, N \to \infty$, so (u_N) is uniformly Cauchy. Hence $u_N \to u$ uniformly on $[0, L]$, and

$$|u_N(x) - u(x)| \le \sum_{m=N+1}^{\infty} |a_m| \longrightarrow 0$$

uniformly in x. Each u_N is continuous, so u is continuous. Finally, $u_N(0) = u_N(L) = 0$ for every N, and uniform convergence gives

$$u(0) = u(L) = 0.$$

□

Chapter 6

Poincaré Inequalities, Compactness, and Weak Elliptic Problems

This chapter develops four tools that are used repeatedly in elliptic partial differential equations: Poincaré inequalities, zero-boundary Sobolev spaces, Rellich compactness, and the Lax–Milgram theorem. Unless stated otherwise, $U \subset \mathbb{R}^m$ is a bounded open set with C^1 boundary.

6.1 Sobolev functions on bounded domains

The global Sobolev inequality on $\mathbb{R}^m$ is transferred to a bounded domain by extending functions across the boundary.

Theorem 6.1 (Smooth approximation up to the boundary). *Let $k \in \mathbb{Z}_{\geq 0}$, $1 \leq p < \infty$, and $u \in W^{k,p}(U)$. There is a sequence $u_n \in C^\infty(\overline{U})$ such that*

$$u_n \longrightarrow u \quad \text{in } W^{k,p}(U).$$

Here $C^\infty(\overline{U})$ denotes restrictions to $\overline{U}$ of functions smooth on a neighborhood of $\overline{U}$.

Theorem 6.2 (Extension theorem). *Let V be any bounded open set with $\overline{U} \subset V$. There is a bounded linear map*

$$E : W^{1,p}(U) \longrightarrow W^{1,p}(\mathbb{R}^m)$$

such that, for every $u \in W^{1,p}(U)$,

$$Eu = u \quad \text{a.e. on } U, \qquad \operatorname{supp}(Eu) \subset V,$$

and

$$\|Eu\|_{W^{1,p}(\mathbb{R}^m)} \leq C \|u\|_{W^{1,p}(U)} .$$

The constant depends on p, U, and V, but not on u.

For first derivatives, the extension theorem and convolution with a mollifier give the preceding approximation theorem; the higher-order argument is analogous. Extension also gives the domain version of the Sobolev estimate.

Theorem 6.3 (Sobolev estimate on a bounded domain)**.** *Suppose* $1 \le p < m$*, and define*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{m}, \qquad p^* = \frac{mp}{m-p}.$$

Then

$$\|u\|_{L^{p^*}(U)} \le C\,\|u\|_{W^{1,p}(U)}$$

for every $u \in W^{1,p}(U)$.

Proof. Extend u to $\widetilde{u} = Eu \in W^{1,p}(\mathbb{R}^m)$ with compact support. Choose $u_n \in C_c^\infty(\mathbb{R}^m)$ with $u_n \to \widetilde{u}$ in $W^{1,p}(\mathbb{R}^m)$. The global Sobolev inequality gives

$$\|u_n - u_\ell\|_{L^{p^*}(\mathbb{R}^m)} \le C\,\|\nabla(u_n - u_\ell)\|_{L^p(\mathbb{R}^m)}\,.$$

Hence (u_n) is Cauchy in L^{p^*}, and passage to the limit yields

$$\|\widetilde{u}\|_{L^{p^*}(\mathbb{R}^m)} \le C\,\|\widetilde{u}\|_{W^{1,p}(\mathbb{R}^m)} \le C\,\|u\|_{W^{1,p}(U)}\,.$$

Restriction to U proves the claim. □

The exponents can be read directly from scaling. If length has physical dimension L, then

$$\|u\|_{L^{p^*}} \sim L^{m/p^*}, \qquad \|\nabla u\|_{L^p} \sim L^{m/p-1}.$$

Thus the homogeneous estimate is scale invariant precisely when

$$\frac{m}{p^*} = \frac{m}{p} - 1. \tag{6.1}$$

On all of $\mathbb{R}^m$, the homogeneous inequality requires compact support or suitable decay. Without such a condition the constant function $u \equiv 1$ is an immediate obstruction.

6.2 Zero boundary values and traces

6.2.1 Boundary values in one and several dimensions

If $f \in H^1(\mathbb{R})$, then f has an absolutely continuous, hence continuous, representative. Indeed,

$$\big|f(x_1) - f(x_2)\big| = \left|\int_{x_2}^{x_1} f'(t)\,\mathrm{d}t\right| \le |x_1 - x_2|^{1/2}\,\big\|f'\big\|_{L^2(\mathbb{R})}\,.$$

In two dimensions, membership in $H^1(\mathbb{R}^2)$ does not assign a meaningful value at an individual point. A codimension-one boundary value does make sense. For $f \in C_c^\infty(\mathbb{R}^2)$,

$$\begin{aligned} f(x_1,0)^2 &= \int_{-\infty}^{0} \partial_2\big(f(x_1,s)^2\big)\,\mathrm{d}s \\ &= 2\int_{-\infty}^{0} f(x_1,s)\,\partial_2 f(x_1,s)\,\mathrm{d}s. \end{aligned}$$

After integration in x_1 and Cauchy–Schwarz,

$$\left\|f(\cdot,0)\right\|_{L^2(\mathbb{R})}^2 \leq 2\left\|f\right\|_{L^2(\mathbb{R}^2)}\left\|\partial_2 f\right\|_{L^2(\mathbb{R}^2)} \leq \left\|f\right\|_{H^1(\mathbb{R}^2)}^2.$$

Density therefore defines the trace $f(\cdot,0)$ for every $f \in H^1(\mathbb{R}^2)$.

Theorem 6.4 (Trace on a coordinate subspace)**.** *Write $(x,y) \in \mathbb{R}^d \times \mathbb{R}^r$. If $s > r/2$, restriction to $y = 0$ extends to a bounded map*

$$H^s(\mathbb{R}^d \times \mathbb{R}^r) \longrightarrow H^{s-r/2}(\mathbb{R}^d), \qquad f \longmapsto f(\cdot,0).$$

More generally, for $1 < p < \infty$ and $s > r/p$, the trace loses r/p derivatives in the $W^{s,p}$ scale:

$$W^{s,p}(\mathbb{R}^d \times \mathbb{R}^r) \longrightarrow W^{s-r/p,p}(\mathbb{R}^d).$$

The second target is understood in the fractional Sobolev–Slobodeckij scale when $s - r/p$ is not an integer.

6.2.2 The space $W_0^{1,p}(U)$

Definition 6.5. The zero-boundary Sobolev space is

$$W_0^{1,p}(U) := \overline{C_c^\infty(U)}^{W^{1,p}(U)}.$$

In particular,

$$H_0^1(U) = W_0^{1,2}(U).$$

More generally, $W_0^{k,p}(U) := \overline{C_c^\infty(U)}^{W^{k,p}(U)}$. On the whole space, compactly supported smooth functions are dense:

$$W_0^{k,p}(\mathbb{R}^m) = W^{k,p}(\mathbb{R}^m).$$

On a bounded domain, $W_0^{k,p}(U)$ is generally a proper subspace of $W^{k,p}(U)$.

For a C^1 domain, this agrees with the kernel of the trace map:

$$W_0^{1,p}(U) = \left\{u \in W^{1,p}(U) : \operatorname{Tr} u = 0 \text{ on } \partial U\right\}.$$

If $u \in W_0^{1,p}(U)$, its extension by zero,

$$\widetilde{u}(x) = \begin{cases} u(x), & x \in U, \\ 0, & x \notin U, \end{cases}$$

belongs to $W^{1,p}(\mathbb{R}^m)$. The weak derivatives do not acquire a boundary measure because the trace of u vanishes.

Theorem 6.6 (Homogeneous estimate on $W_0^{1,p}$)**.** *Let $1 \leq p < m$. For every $1 \leq q \leq p^*$,*

$$\|u\|_{L^q(U)} \leq C\,\|\nabla u\|_{L^p(U)}, \qquad u \in W_0^{1,p}(U).$$

Proof. Take $u_n \in C_c^\infty(U)$ with $u_n \to u$ in $W^{1,p}(U)$, and extend every u_n by zero. The global Sobolev inequality gives

$$\|u_n\|_{L^{p^*}(U)} \le C \|\nabla u_n\|_{L^p(U)} .$$

Passing to the limit proves the endpoint $q = p^*$. Since $|U| < \infty$, Hölder's inequality gives every $1 \le q \le p^*$. □

In particular,

$$\|u\|_{W^{1,p}(U)} \asymp \|\nabla u\|_{L^p(U)} \quad \text{on } W_0^{1,p}(U).$$

For $p = 2$, the right-hand side defines the equivalent Hilbert norm used for the Dirichlet problem below.

6.3 Compact Sobolev embeddings

The Sobolev embedding at the critical exponent is continuous but is not compact. Compactness is recovered strictly below p^*.

Definition 6.7. Let X and Y be Banach spaces with $X \subset Y$. The notation $X \hookrightarrow Y$ means that the inclusion is continuous. The notation $X \hookrightarrow\hookrightarrow Y$ means that the inclusion is compact: every bounded sequence in X has a subsequence converging in Y.

Theorem 6.8 (Rellich compactness)**.** *Suppose $1 \le p < m$. Then*

$$W^{1,p}(U) \hookrightarrow\hookrightarrow L^q(U) \qquad \textit{for every } 1 \le q < p^*.$$

Proof. Let (u_n) be bounded in $W^{1,p}(U)$. By the extension theorem, we may regard every u_n as a function on $\mathbb{R}^m$, with all supports contained in one bounded open set V, and with

$$\sup_n \|u_n\|_{W^{1,p}(\mathbb{R}^m)} < \infty.$$

Let $\eta \in C_c^\infty(B_1(0))$ be a standard mollifier, set $\eta_\varepsilon(x) = \varepsilon^{-m}\eta(x/\varepsilon)$, and write $u_n^\varepsilon = \eta_\varepsilon * u_n$.

The fundamental theorem of calculus along line segments gives

$$\begin{aligned} u_n^\varepsilon(x) - u_n(x) &= \int_{B_1(0)} \eta(z)\big(u_n(x - \varepsilon z) - u_n(x)\big)\, \mathrm{d}z \\ &= -\varepsilon \int_{B_1(0)} \eta(z) \int_0^1 \nabla u_n(x - t\varepsilon z) \cdot z \, \mathrm{d}t \, \mathrm{d}z. \end{aligned}$$

Fubini, translation invariance, and Hölder's inequality on the fixed set V imply

$$\|u_n^\varepsilon - u_n\|_{L^1(\mathbb{R}^m)} \le C\varepsilon \|\nabla u_n\|_{L^1(\mathbb{R}^m)} \le C\varepsilon.$$

The Sobolev inequality gives a uniform L^{p^*} bound. Fix $1 \le q < p^*$, and choose $0 < \theta \le 1$ so that

$$\frac{1}{q} = \theta + \frac{1-\theta}{p^*}.$$

Interpolation yields

$$\|u_n^\varepsilon - u_n\|_{L^q} \le \|u_n^\varepsilon - u_n\|_{L^1}^{\theta} \, \|u_n^\varepsilon - u_n\|_{L^{p^*}}^{1-\theta} \le C\varepsilon^{\theta},$$

uniformly in n.

This smoothing estimate is the averaged form of the translation bound used in the lecture. For a Sobolev function on $\mathbb{R}^m$,

$$\int_{\mathbb{R}^m} |u(x+h) - u(x)|^p \, \mathrm{d}x \le |h|^p \int_{\mathbb{R}^m} |\nabla u(x)|^p \, \mathrm{d}x.$$

It follows from $u(x+h) - u(x) = \int_0^1 \nabla u(x+th) \cdot h \, \mathrm{d}t$ and Hölder's inequality.

For fixed $\varepsilon > 0$, the family (u_n^ε) is uniformly bounded and equicontinuous. Indeed,

$$|u_n^\varepsilon(x)| \le \left\|\eta_\varepsilon\right\|_{L^\infty} \|u_n\|_{L^1} \le C_\varepsilon,$$

and

$$\left|u_n^\varepsilon(x) - u_n^\varepsilon(y)\right| \le \|\nabla u_n^\varepsilon\|_{L^\infty} \left|x - y\right| \le C_\varepsilon \left|x - y\right|.$$

Their supports lie in a fixed bounded set. Arzelà–Ascoli therefore produces a uniformly convergent subsequence for every fixed ε.

Choose a sequence $\varepsilon_k \downarrow 0$ and apply a diagonal argument. The resulting subsequence, still denoted (u_n), is Cauchy in L^q: for fixed k,

$$\left\|u_j - u_\ell\right\|_{L^q} \le \left\|u_j - u_j^{\varepsilon_k}\right\|_{L^q} + \left\|u_j^{\varepsilon_k} - u_\ell^{\varepsilon_k}\right\|_{L^q} + \left\|u_\ell^{\varepsilon_k} - u_\ell\right\|_{L^q}.$$

First choose k so that the first and third terms are small, and then choose j, ℓ large so that the middle term is small. Thus the subsequence converges in $L^q(U)$. □

Remark 6.9. The embedding $W^{1,p}(U) \hookrightarrow L^{p^*}(U)$ is generally not compact. On a bounded C^1 domain, however, $W^{1,p}(U) \hookrightarrow\hookrightarrow L^p(U)$ holds for every $1 \le p < \infty$. For $p > m$, this follows from the Hölder regularity supplied by Morrey's estimate; for $p = \infty$, bounded sets in $W^{1,\infty}$ are uniformly bounded and equi-Lipschitz.

6.4 Poincaré inequalities

For $u \in L^1(U)$, write

$$u_U := \fint_U u(y) \, \mathrm{d}y = \frac{1}{|U|} \int_U u(y) \, \mathrm{d}y.$$

Theorem 6.10 (Poincaré inequality)**.** *Suppose U is connected. For every $1 \le p \le \infty$, there is a constant $C = C(U, p)$ such that*

$$\|u - u_U\|_{L^p(U)} \le C \, \|\nabla u\|_{L^p(U)}$$

for every $u \in W^{1,p}(U)$.

Proof. We first take $1 \le p < m$, which is the range used in the compactness argument above. If the estimate were false, there would be $u_n \in W^{1,p}(U)$ such that

$$\|u_n - (u_n)_U\|_{L^p(U)} > n \|\nabla u_n\|_{L^p(U)} .$$

Normalize by setting

$$v_n = \frac{u_n - (u_n)_U}{\|u_n - (u_n)_U\|_{L^p(U)}}.$$

Then

$$\|v_n\|_{L^p(U)} = 1, \qquad (v_n)_U = 0, \qquad \|\nabla v_n\|_{L^p(U)} < \frac{1}{n}.$$

Thus (v_n) is bounded in $W^{1,p}(U)$. Rellich compactness gives a subsequence converging in $L^p(U)$ to some v. For every $\varphi \in C_c^\infty(U)$,

$$\left\langle v, \partial_i \varphi \right\rangle = \lim_{n\to\infty} \left\langle v_n, \partial_i \varphi \right\rangle = -\lim_{n\to\infty} \left\langle \partial_i v_n, \varphi \right\rangle = 0.$$

Hence $\partial_i v = 0$ weakly for every i. Connectedness gives that v is constant. Since $v_U = 0$, we have $v = 0$, contradicting $\|v\|_{L^p(U)} = 1$.

The remaining finite exponents follow from the corresponding compact embedding $W^{1,p}(U) \hookrightarrow\hookrightarrow L^p(U)$, with the same contradiction argument. If $p = \infty$, extend u to $Eu \in W^{1,\infty}(\mathbb{R}^m)$. Its Lipschitz representative satisfies

$$\left|u(x) - u(y)\right| \le C \left|x - y\right| \|\nabla u\|_{L^\infty(U)} \qquad (x, y \in U).$$

Averaging in y gives the asserted L^∞ estimate. □

This gives a useful Hilbert space for the Neumann problem:

$$\widetilde{H}^1(U) := \left\{u \in H^1(U) : u_U = 0\right\}.$$

On $\widetilde{H}^1(U)$,

$$\|u\|_{\widetilde{H}^1} := \|\nabla u\|_{L^2(U)}$$

is a norm equivalent to the ordinary H^1 norm.

6.4.1 The scale of the inequality on balls

For a ball $B_r(x_0)$, scaling the unit-ball estimate gives

$$\left\|u - u_{B_r(x_0)}\right\|_{L^p(B_r(x_0))} \le C(m,p) r \|\nabla u\|_{L^p(B_r(x_0))} . \tag{6.2}$$

Indeed, set $v(y) = u(x_0 + ry)$ on $B_1(0)$, apply the unit-ball inequality, and change variables back to x.

The line-segment calculation behind this estimate is

$$u(x_0 + \rho\omega) - u(x_0) = \int_0^\rho \partial_r u(x_0 + s\omega)\, \mathrm{d}s,$$

followed by integration in polar coordinates. The geometry is the one shown below.

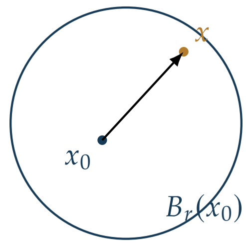


Figure 6.1: Integration along a radial segment in a ball.

Combining Poincaré with the Sobolev estimate yields the Poincaré–Sobolev form

$$\left\|u - u_{B_r(x_0)}\right\|_{L^{p^*}(B_r(x_0))} \le C(m,p)\, \|\nabla u\|_{L^p(B_r(x_0))}, \qquad 1 \le p < m.$$

The absence of r here agrees with the dimensional identity (6.1).

6.5 The Lax–Milgram theorem

Let H be a real Hilbert space. A bilinear form $B : H \times H \to \mathbb{R}$ is *bounded* if

$$|B(u,v)| \le M\, \|u\|_H\, \|v\|_H,$$

and *coercive* if

$$B(u,u) \ge c\, \|u\|_H^2$$

for constants $M < \infty$ and $c > 0$.

Theorem 6.11 (Lax–Milgram). *Let B be a bounded, coercive bilinear form on H. For every $F \in H'$, there is a unique $u \in H$ such that*

$$B(u,v) = F(v) \qquad \textit{for every } v \in H.$$

Moreover,

$$\|u\|_H \le c^{-1}\, \|F\|_{H'}.$$

Proof. By the Riesz representation theorem, there is a bounded operator $A : H \to H$ such that

$$B(u,v) = \langle Au, v\rangle_H.$$

Coercivity and Cauchy–Schwarz give

$$c\, \|u\|_H^2 \le \langle Au, u\rangle_H \le \|Au\|_H\, \|u\|_H,$$

so

$$\|Au\|_H \ge c\, \|u\|_H.$$

Thus A is injective and its range is closed. If $y \perp \operatorname{Ran} A$, then $0 = \left\langle Ax, y\right\rangle_H = B(x,y)$ for every $x \in H$. Taking $x = y$ and using coercivity gives $y = 0$. Hence the range is dense as well as closed, and therefore A is onto. The equation $B(u,v) = F(v)$ is now the Riesz form of $Au = f$, which has a unique solution. The same coercive estimate gives the stated bound. □

6.6 Weak solutions of elliptic boundary-value problems

Elliptic regularity asks how much additional smoothness a weak solution inherits from the right-hand side, the coefficients, and the boundary. Before addressing that question, one first needs a weak solution. The existence theory in this lecture comes from Lax–Milgram.

6.6.1 The Dirichlet problem

Consider

$$\begin{cases} -\Delta u = f & \text{in } U, \\ u = 0 & \text{on } \partial U, \end{cases} \tag{6.3}$$

with $f \in L^2(U)$. Multiplication by $v \in H_0^1(U)$ and integration by parts lead to

$$\int_U \nabla u \cdot \nabla v \, \mathrm{d}x = \int_U f v \, \mathrm{d}x.$$

This identity is taken as the definition of a weak solution. Set

$$H = H_0^1(U), \qquad B(u,v) = \int_U \nabla u \cdot \nabla v \, \mathrm{d}x, \qquad F(v) = \int_U f v \, \mathrm{d}x.$$

The equivalent norm on $H_0^1(U)$, Poincaré's inequality, and Cauchy–Schwarz show that B is bounded and coercive and that $F \in H'$. Lax–Milgram therefore gives a unique weak solution $u \in H_0^1(U)$.

For inhomogeneous data $u = g$ on ∂U, choose $\phi \in H^2(U)$ with trace g, and set $w = u - \phi$. Then

$$w \in H_0^1(U), \qquad -\Delta w = f + \Delta \phi.$$

The homogeneous theory applies to w.

6.6.2 A uniformly elliptic divergence-form operator

Let

$$Lu = -\sum_{i,j=1}^{m} \partial_i \big(a^{ij}(x) \partial_j u\big),$$

where $a^{ij} = a^{ji}$ are smooth and satisfy the uniform ellipticity condition

$$\sum_{i,j=1}^{m} a^{ij}(x) \xi_i \xi_j \geq \lambda \, |\xi|^2 \quad \text{for every } x \in U, \ \xi \in \mathbb{R}^m, \tag{6.4}$$

for some $\lambda > 0$. The quadratic expression in (6.4) is the characteristic form of the principal part. Its positivity for every $\xi \neq 0$ is precisely ellipticity.

The weak Dirichlet form is

$$B(u,v) = \int_U \sum_{i,j=1}^{m} a^{ij}(x) \partial_i u \, \partial_j v \, \mathrm{d}x.$$

Boundedness of the coefficients gives

$$|B(u,v)| \leq C \, \|\nabla u\|_{L^2(U)} \, \|\nabla v\|_{L^2(U)} \, ,$$

while uniform ellipticity gives

$$B(u,u) \geq \lambda \, \|\nabla u\|^2_{L^2(U)} \, .$$

Thus Lax–Milgram yields a unique weak solution in $H^1_0(U)$ for every right-hand side in $H^{-1}(U) = (H^1_0(U))'$, in particular for every $f \in L^2(U)$.

6.6.3 The Neumann problem

Now consider

$$\begin{cases} -\Delta u = f & \text{in } U, \\ \partial_\nu u = 0 & \text{on } \partial U, \end{cases} \tag{6.5}$$

where ν is the outward unit normal. The derivative boundary condition is natural in the weak formulation, so the appropriate space is $H^1(U)$, not $H^1_0(U)$. Integration by parts gives

$$\int_U \nabla u \cdot \nabla v \, \mathrm{d}x = \int_U f v \, \mathrm{d}x \qquad \text{for every } v \in H^1(U).$$

There are two obstructions to applying Lax–Milgram directly on $H^1(U)$. First, constants lie in the kernel of the form

$$B(u,v) = \int_U \nabla u \cdot \nabla v \, \mathrm{d}x,$$

so coercivity fails and solutions cannot be unique. Second, choosing $v \equiv 1$ in the weak equation gives the necessary compatibility condition

$$\int_U f \, \mathrm{d}x = 0. \tag{6.6}$$

Assume (6.6). Work on the mean-zero Hilbert space $\widetilde{H}^1(U)$. Poincaré's inequality makes

$$\|u\|_{\widetilde{H}^1} = \|\nabla u\|_{L^2(U)}$$

equivalent to the H^1 norm, and B is coercive in this norm. The functional $F(v) = \int_U f v \, \mathrm{d}x$ is bounded. Lax–Milgram therefore gives a unique mean-zero weak solution. Every other weak solution differs from it by a constant.

Finally, suppose that this weak solution is smooth. Testing first with $v \in C^\infty_c(U)$ recovers $-\Delta u = f$ in the interior. Green's identity then gives, for arbitrary smooth v,

$$\int_{\partial U} (\partial_\nu u) v \, \mathrm{d}S - \int_U (\Delta u) v \, \mathrm{d}x = \int_U \nabla u \cdot \nabla v \, \mathrm{d}x = \int_U f v \, \mathrm{d}x.$$

Since $-\Delta u = f$, the interior terms cancel and

$$\int_{\partial U} (\partial_\nu u) v \, \mathrm{d}S = 0$$

for every boundary trace v. Hence $\partial_\nu u = 0$ on ∂U, showing how the Neumann boundary condition is encoded by the weak formulation.

Chapter 7

Caccioppoli Estimates and Elliptic Boundary Regularity

The basic existence theory for weak elliptic equations gives solutions in $H_0^1(\Omega)$. The aim of this chapter is to recover the additional derivatives hidden in the equation. We first prove a local energy estimate, then use difference quotients to obtain interior regularity, and finally extend the argument to the boundary. The same ideas also apply to the Neumann problem. These are a priori arguments: they establish properties of a solution once it exists, while potential theory provides a separate route to constructing solutions.

Throughout, $\Omega \subset \mathbb{R}^m$ is a bounded, connected domain unless stated otherwise.

7.1 Weak equations and the natural test space

Consider the homogeneous Dirichlet problem

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{7.1}$$

where $f \in L^2(\Omega)$. Its weak formulation is

$$\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x \qquad \text{for every } v \in H_0^1(\Omega). \tag{7.2}$$

Riesz representation, or equivalently the Lax–Milgram theorem in this case, gives a unique $u \in H_0^1(\Omega)$.

The space

$$H^{-1}(\Omega) := \big(H_0^1(\Omega)\big)^*$$

is the natural range of the weak Laplacian. Indeed, if $u \in H_0^1(\Omega)$, then

$$\langle -\Delta u, v\rangle := \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x$$

defines an element of $H^{-1}(\Omega)$. Equivalently, the extension of u by zero belongs to $H^1(\mathbb{R}^m)$, and the Fourier-transform description of Sobolev spaces gives

$$\Delta : H^1(\mathbb{R}^m) \longrightarrow H^{-1}(\mathbb{R}^m).$$

There is an important test-space distinction. A distribution in $H^{-1}(\Omega)$ acts automatically on $H^1_0(\Omega)$, but not on all of $H^1(\Omega)$. Thus a weak equation with test functions that need not vanish on the boundary carries extra boundary information and must be formulated separately.

The same weak framework applies to a divergence-form equation

$$-\partial_i\big(a^{ij}(x)\partial_j u\big) = f. \tag{7.3}$$

If the matrix $A = (a^{ij})$ is bounded and uniformly positive definite, then the corresponding identity is

$$\int_\Omega a^{ij}\partial_j u\,\partial_i v\,\mathrm{d}x = \int_\Omega f v\,\mathrm{d}x.$$

7.2 The Caccioppoli inequality

Let $B_r(x_0) \Subset B_R(x_0) \Subset \Omega$. Choose a cutoff $\eta \in C_c^\infty(B_R(x_0))$ satisfying

$$0 \le \eta \le 1, \qquad \eta = 1 \text{ on } B_r(x_0), \qquad |\nabla\eta| \le \frac{C}{R-r}. \tag{7.4}$$

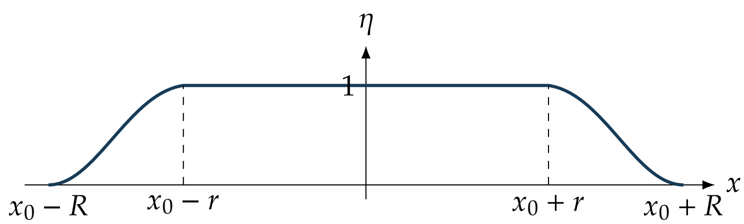


Figure 7.1: A cutoff equal to one on the smaller ball and supported in the larger ball.

Theorem 7.1 (Caccioppoli inequality)**.** *Suppose $u \in H^1(\Omega)$ is a weak solution of $-\Delta u = f$ with $f \in L^2(\Omega)$. Then*

$$\int_{B_r(x_0)} |\nabla u|^2\,\mathrm{d}x \le C\left(1 + \frac{1}{(R-r)^2}\right)\int_{B_R(x_0)} |u|^2\,\mathrm{d}x + C\int_{B_R(x_0)} |f|^2\,\mathrm{d}x. \tag{7.5}$$

The same conclusion holds for (7.3), with a constant depending also on the ellipticity bounds of A.

Proof. The function $v = \eta^2 u$ is an admissible test function. Substitution into (7.2) gives

$$\int_\Omega \eta^2|\nabla u|^2\,\mathrm{d}x + 2\int_\Omega \eta u\,\nabla\eta\cdot\nabla u\,\mathrm{d}x = \int_\Omega f\eta^2 u\,\mathrm{d}x. \tag{7.6}$$

Hence, by Cauchy–Schwarz and Young's inequality,

$$\begin{aligned}\|\eta\nabla u\|_{L^2}^2 &\le 2\|\eta\nabla u\|_{L^2}\|u\nabla\eta\|_{L^2} + \|\eta f\|_{L^2}\|\eta u\|_{L^2}\\ &\le \frac{1}{2}\|\eta\nabla u\|_{L^2}^2 + C\|u\nabla\eta\|_{L^2}^2 + C\|\eta f\|_{L^2}^2 + C\|\eta u\|_{L^2}^2.\end{aligned}$$

Absorb the first term on the right, use (7.4), and recall that $\eta = 1$ on $B_r(x_0)$. This proves (7.5).

For a uniformly elliptic coefficient matrix, the left side of (7.6) is bounded below by a fixed multiple of $\|\eta\nabla u\|_{L^2}^2$; the remaining terms are estimated in exactly the same way. □

The value of the estimate is that it controls one derivative of u on a smaller ball using only u and the equation on a larger ball. Applying this idea to difference quotients will produce a second derivative.

7.3 Interior regularity by difference quotients

Suppose

$$u \in H^1_0(\Omega), \qquad -\Delta u = f \in L^2(\Omega). \tag{7.7}$$

A tempting formal argument is to differentiate the equation, multiply by a cutoff times a second derivative of u, and integrate by parts. This is not yet legitimate: the second derivative is precisely what we are trying to prove exists, and an expression such as $\partial_{ij}(\eta^2 u)$ need not belong to $H^1_0(\Omega)$. The useful estimate behind the formal calculation is to bound the highest-order term by a small multiple of itself plus lower-order terms and then absorb the small multiple. The difference-quotient argument below makes that procedure legitimate.

One repair is to mollify u, carry out the calculation for $u_\varepsilon = u * \rho_\varepsilon$ away from the boundary, and pass to the limit. Difference quotients give a direct weak argument.

For $1 \le k \le m$ and $h \neq 0$, define

$$\delta_{h,k} u(x) := \frac{u(x + he_k) - u(x)}{h}. \tag{7.8}$$

If η is supported strictly inside Ω and $|h|$ is small enough, then

$$v = -\delta_{-h,k}\left(\eta^2 \delta_{h,k} u\right) \tag{7.9}$$

belongs to $H^1_0(\Omega)$ and is therefore an admissible test function.

The discrete integration-by-parts identity

$$\int g\, \delta_{-h,k} w \, \mathrm{d}x = -\int \delta_{h,k} g \, w \, \mathrm{d}x$$

transforms the left side of the weak equation into

$$\begin{aligned}
\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x &= \int_\Omega \delta_{h,k}(\partial_i u)\, \partial_i\left(\eta^2 \delta_{h,k} u\right) \mathrm{d}x \\
&= \int_\Omega \eta^2 \left|\nabla \delta_{h,k} u\right|^2 \, \mathrm{d}x + 2\int_\Omega \eta\, \delta_{h,k}(\partial_i u)\, \partial_i \eta\, \delta_{h,k} u \, \mathrm{d}x.
\end{aligned}$$

For the right side, the elementary difference-quotient estimate

$$\left\|\delta_{-h,k} w\right\|_{L^2} \le \|\partial_k w\|_{L^2} \qquad (w \in H^1_0)$$

allows the terms containing $\nabla \delta_{h,k} u$ to be absorbed into the left side. Young's inequality then yields the uniform estimate

$$\int_\Omega \eta^2 \left|\nabla \delta_{h,k} u\right|^2 \, \mathrm{d}x \le C \int_{\operatorname{supp} \eta} |f|^2 \, \mathrm{d}x + C_\eta \int_{\operatorname{supp} \eta} \left|\delta_{h,k} u\right|^2 \, \mathrm{d}x. \tag{7.10}$$

The last integral is uniformly bounded because

$$\left\|\delta_{h,k}u\right\|_{L^2} \le \|\partial_k u\|_{L^2}\,.$$

Applying the first Caccioppoli estimate to u controls this remaining first derivative by u and f.

The difference-quotient characterization of H^1, followed by weak compactness as $h \to 0$, now gives $\partial_{ik}u \in L^2_{\mathrm{loc}}(\Omega)$ for every i, k. Thus we obtain the following result.

Theorem 7.2 (Interior H^2 regularity)**.** *If $u \in H^1(\Omega)$ is a weak solution of $-\Delta u = f$ and $f \in L^2(\Omega)$, then*

$$u \in H^2_{\mathrm{loc}}(\Omega).$$

More precisely, whenever $B_r(x_0) \Subset B_R(x_0) \Subset \Omega$,

$$\sum_{i,j=1}^{m} \int_{B_r(x_0)} \left|\partial_{ij}u\right|^2 \,\mathrm{d}x \le C_{r,R}\left(\int_{B_R(x_0)} \left|f\right|^2 \,\mathrm{d}x + \int_{B_R(x_0)} |u|^2 \,\mathrm{d}x\right). \tag{7.11}$$

The same argument can be iterated. If $f \in H^k_{\mathrm{loc}}(\Omega)$, then

$$u \in H^{k+2}_{\mathrm{loc}}(\Omega). \tag{7.12}$$

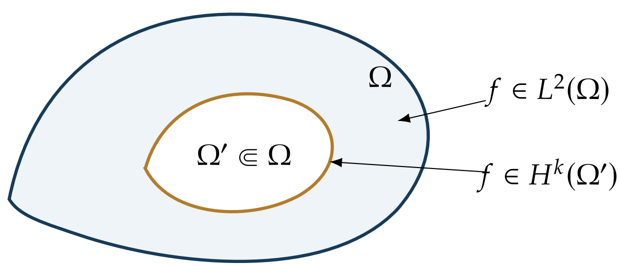


Figure 7.2: Higher regularity is local: extra smoothness of f on an interior subdomain gives extra smoothness of u there.

At each stage, difference quotients move one derivative onto the equation, and a localized Caccioppoli estimate controls the next derivative of u. The shrinking from B_R to B_r is essential: translations in a difference quotient cannot be used uniformly all the way to a curved boundary.

7.4 Regularity up to the boundary

7.4.1 The flat-boundary argument

First suppose that a portion of the boundary is flat. In local coordinates write

$$B_R^+ = B_R \cap \{x_m > 0\}\,, \qquad \Gamma_R = B_R \cap \{x_m = 0\}\,,$$

and suppose $u = 0$ on Γ_R. A translation in a tangential direction e_α, $1 \le \alpha < m$, respects the flat boundary. The difference-quotient argument therefore controls

$$\partial_{i\alpha} u \in L^2(B_r^+) \qquad (1 \le i \le m,\ 1 \le \alpha < m).$$

In particular, all tangential–tangential and normal–tangential second derivatives are controlled. The equation supplies the missing pure normal derivative:

$$\partial_{mm} u = -f - \sum_{\alpha=1}^{m-1} \partial_{\alpha\alpha} u. \tag{7.13}$$

Hence $u \in H^2$ up to the flat part of the boundary.

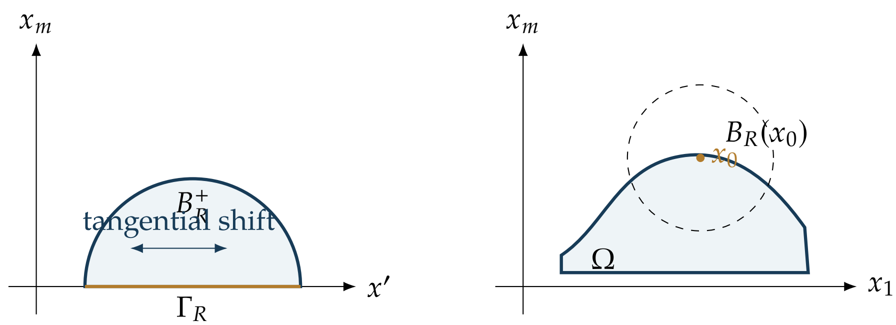


Figure 7.3: A flat boundary permits tangential translations. Near a curved boundary, one first localizes and then straightens the boundary.

7.4.2 Flattening a curved boundary

Near a point $x_0 \in \partial\Omega$, take a partition of unity, localize to a small coordinate neighborhood, and choose a C^2 diffeomorphism

$$y = \Phi(x), \qquad x = \Psi(y),$$

that sends the curved boundary to $\{y_m = 0\}$. If $\widetilde{u}(y) = u(\Psi(y))$, the chain rule gives

$$\partial_i u(x) = \partial_\alpha \widetilde{u}(y)\, \partial_i \Phi^\alpha(x). \tag{7.14}$$

Consequently, the Dirichlet energy transforms into a divergence-form bilinear form:

$$\int a^{\alpha\beta}(y) \partial_\beta \widetilde{u}\, \partial_\alpha \widetilde{v}\, \mathrm{d}y, \qquad a^{\alpha\beta}(y) = J(y) \sum_{i=1}^{m} \partial_i \Phi^\alpha(\Psi(y)) \partial_i \Phi^\beta(\Psi(y)), \tag{7.15}$$

where $J(y) = \left|\det D\Psi(y)\right|$. The right side becomes $\widetilde{f}(y) = J(y) f(\Psi(y))$. The matrix in (7.15) is uniformly elliptic on a sufficiently small coordinate patch. Thus the flat-boundary argument for a general uniformly elliptic coefficient matrix applies, and a finite partition of unity gives a global estimate.

Theorem 7.3 (Global Dirichlet regularity). *Let $k \in \mathbb{Z}_{\geq 0}$, let Ω be a bounded domain with boundary of class C^{k+2}, and let $f \in H^k(\Omega)$. If $u \in H_0^1(\Omega)$ is the weak solution of*

$$-\Delta u = f,$$

then

$$u \in H_0^1(\Omega) \cap H^{k+2}(\Omega)$$

and

$$\|u\|_{H^{k+2}(\Omega)} \leq C \left(\|f\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)} \right). \tag{7.16}$$

The boundary-flattening calculation naturally leads to the corresponding variable-coefficient statement.

Theorem 7.4 (Divergence-form regularity). *Let $k \in \mathbb{Z}_{\geq 0}$, suppose $a^{ij} \in W^{k+1,\infty}(\Omega)$, and assume that $A = (a^{ij})$ is symmetric and uniformly elliptic: there are constants $0 < c \leq C < \infty$ such that*

$$c\,|\xi|^2 \leq a^{ij}(x)\xi_i\xi_j \leq C\,|\xi|^2 \qquad (x \in \Omega,\ \xi \in \mathbb{R}^m). \tag{7.17}$$

If $\partial\Omega$ is of class C^{k+2}, $f \in H^k(\Omega)$, and $u \in H_0^1(\Omega)$ solves

$$-\partial_i\left(a^{ij}\partial_j u\right) = f,$$

then $u \in H_0^1(\Omega) \cap H^{k+2}(\Omega)$.

7.5 The homogeneous Neumann problem

Consider

$$-\Delta u = f \quad \text{in } \Omega, \qquad \partial_\nu u = 0 \quad \text{on } \partial\Omega. \tag{7.18}$$

Integrating the equation shows that the compatibility condition

$$\int_\Omega f \,\mathrm{d}x = 0 \tag{7.19}$$

is necessary. Constants lie in the kernel of the Neumann Laplacian, so a normalization is also required. Write

$$u_\Omega := \frac{1}{|\Omega|}\int_\Omega u \,\mathrm{d}x, \qquad \widetilde{H}^1(\Omega) := \left\{ v \in H^1(\Omega) : v_\Omega = 0 \right\}.$$

The Poincaré inequality makes $\|\nabla v\|_{L^2}$ an equivalent norm on $\widetilde{H}^1(\Omega)$. Hence Lax–Milgram gives a unique normalized solution $u \in \widetilde{H}^1(\Omega)$ satisfying

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega f v \,\mathrm{d}x \qquad \text{for every } v \in \widetilde{H}^1(\Omega). \tag{7.20}$$

The local Caccioppoli argument is unchanged after replacing u by $u - u_\Omega$. More precisely, start with $w = \eta^2(u - u_\Omega)$ and use the admissible mean-zero test function $v = w - w_\Omega$. Constants disappear from the gradient, and (7.19) removes the constant from the right side. Difference quotients then give

$$u \in \widetilde{H}^1(\Omega) \cap H^2_{\mathrm{loc}}(\Omega).$$

The flat-boundary proof did not rely on a zero trace when it estimated tangential difference quotients; for the Neumann formulation the same tests are admissible in H^1. After flattening the boundary, the same argument therefore gives $u \in H^2(\Omega)$ when the boundary is sufficiently smooth.

Once this regularity is known, integration by parts makes the boundary condition explicit. Since the compatibility condition extends (7.20) from mean-zero tests to all $v \in H^1(\Omega)$,

$$\begin{aligned}\int_\Omega f v \, \mathrm{d}x &= \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x \\ &= -\int_\Omega \Delta u \, v \, \mathrm{d}x + \int_{\partial\Omega} \partial_\nu u \, v \, \mathrm{d}S.\end{aligned}$$

Because $-\Delta u = f$, it follows that

$$\int_{\partial\Omega} \partial_\nu u \, v \, \mathrm{d}S = 0 \qquad \text{for every } v \in H^1(\Omega),$$

and hence $\partial_\nu u = 0$ on $\partial\Omega$ in the trace sense. Thus the weak construction indeed solves the Neumann problem, uniquely up to an additive constant and uniquely in $\widetilde{H}^1(\Omega)$.

7.6 Review checklist

The lecture closes with the following list of themes to organize the elliptic theory developed so far.

- Know the principal model partial differential equations and which ones can be solved explicitly or abstractly. For each problem, separate existence from uniqueness. Relevant tools include the Cauchy–Kowalevskaya theorem and the maximum principle for harmonic functions.
- Be comfortable with distributions and the Fourier transform.
- Know the elliptic regularity theory, including C^α regularity and the estimates for singular integral operators.
- Recognize when a problem is elliptic and understand the role of characteristic directions.
- Be able to state precisely which interior and boundary regularity results have been established and which hypotheses each result requires.

Chapter 8

Boundary Conditions, Higher-Order Elliptic Problems, and the Fredholm Alternative

Throughout this chapter, $\Omega \subset \mathbb{R}^m$ is a bounded domain. Whenever boundary traces or elliptic regularity are used, the boundary and the coefficients are assumed to be sufficiently regular for the stated operation. Repeated indices are summed.

8.1 From a coercive form back to the boundary-value problem

Consider the bilinear form

$$D(u, v) = \int_\Omega \left(a_{ij}\, \partial_i u\, \partial_j v + b_i\, \partial_i u\, v + c_i\, \partial_i v\, u + d\, uv\right)\, \mathrm{d}x. \tag{8.1}$$

Let X be the Hilbert space in which the weak problem is to be solved. The Lax–Milgram argument requires constants $C_1, C_2 > 0$ such that

$$D(u, u) \geq C_1 \|u\|_X^2, \tag{8.2}$$
$$|D(u, v)| \leq C_2 \|u\|_X \|v\|_X. \tag{8.3}$$

Under the usual boundedness assumptions on the coefficients, (8.3) follows from Cauchy–Schwarz and Hölder whenever X is continuously embedded in $H^1(\Omega)$. If $f \in L^2(\Omega)$, then

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

is a bounded linear functional on such a space. Lax–Milgram therefore gives a unique $u \in X$ satisfying

$$D(u, v) = F(v) \qquad \text{for every } v \in X. \tag{8.4}$$

The choice of X contains the *essential* boundary conditions. The remaining, or *natural*, boundary conditions are encoded by the form. To see them, suppose elliptic regularity gives enough regularity to integrate (8.1) by parts. Then

$$\begin{aligned} D(u,v) = &\int_\Omega \left[-\partial_j\big(a_{ij}\partial_i u\big) + b_i\partial_i u - \partial_i(c_i u) + du\right] v \, \mathrm{d}x \\ &+ \int_{\partial\Omega} \big(a_{ij}\nu_j\partial_i u + c_i\nu_i u\big)\, v \, \mathrm{d}S, \end{aligned} \tag{8.5}$$

where ν is the outward unit normal. Thus the associated differential operator is

$$Lu = -\partial_j\big(a_{ij}\partial_i u\big) + b_i\partial_i u - \partial_i(c_i u) + du. \tag{8.6}$$

If $C_c^\infty(\Omega) \subset X$, then testing with compactly supported functions gives

$$Lu = f$$

in the sense of distributions, and almost everywhere whenever both sides are represented by functions. Once the interior equation is known, (8.5) leaves

$$\int_{\partial\Omega} \big(a_{ij}\nu_j\partial_i u + c_i\nu_i u\big)\, v \, \mathrm{d}S = 0 \qquad (v \in X). \tag{8.7}$$

If the traces of functions in X are arbitrary enough, this identity yields the natural boundary condition

$$a_{ij}\nu_j\partial_i u + c_i\nu_i u = 0 \qquad \text{on } \partial\Omega. \tag{8.8}$$

The role of regularity

The weak equation always makes sense at the dual-space level. If merely $f \in H^{-1}(\Omega)$, one generally obtains $Lu = f$ only as an identity in $H^{-1}(\Omega)$. Without additional elliptic regularity, there may be too little regularity to integrate by parts again and identify the boundary condition pointwise. Boundary conditions must then remain encoded in the choice of X and in the weak form.

8.2 Fourth-order equations and boundary traces

For a second-order problem such as

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ \partial_\nu u = 0 & \text{on } \partial\Omega, \end{cases} \tag{8.9}$$

the energy space is a subspace of $H^1(\Omega)$. The normal derivative is not imposed through the definition of $H^1(\Omega)$; it emerges as a natural boundary condition.

The fourth-order analogue is the biharmonic equation

$$\Delta^2 u = f. \tag{8.10}$$

Its natural energy space is a subspace of $H^2(\Omega)$. Formally, a Cauchy problem for a fourth-order equation prescribes normal derivatives of orders $0, 1, 2,$ and 3 on a noncharacteristic hypersurface.

A boundary-value problem is different: the trace theorem limits which conditions can be built directly into the H^2 energy space.

For $u \in H^2(\Omega)$, the traces satisfy

$$u\big|_{\partial\Omega} \in H^{3/2}(\partial\Omega), \qquad \partial_i u\big|_{\partial\Omega} \in H^{1/2}(\partial\Omega) \quad (1 \le i \le m).$$

In particular, $\partial_\nu u|_{\partial\Omega} \in H^{1/2}(\partial\Omega)$. Second derivatives of a general H^2 function do not possess an ordinary boundary trace. Consequently, zeroth- and first-order conditions can be essential conditions in H^2, whereas a condition involving Δu must normally be recovered from the weak equation after regularity is known.

8.2.1 Clamped and Navier boundary conditions

For the clamped problem

$$\begin{cases} \Delta^2 u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \\ \partial_\nu u = 0 & \text{on } \partial\Omega, \end{cases} \tag{8.11}$$

both boundary conditions can be incorporated in

$$X = H_0^2(\Omega) = \left\{ u \in H^2(\Omega) : u = \partial_\nu u = 0 \text{ on } \partial\Omega \right\},$$

where the trace characterization is understood for a sufficiently regular boundary.

Now consider the Navier, or simply supported, problem

$$\begin{cases} \Delta^2 u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \\ \Delta u = 0 & \text{on } \partial\Omega. \end{cases} \tag{8.12}$$

Only the first boundary condition can be placed directly in the energy space:

$$X = H^2(\Omega) \cap H_0^1(\Omega) = \left\{ u \in H^2(\Omega) : u = 0 \text{ on } \partial\Omega \right\}. \tag{8.13}$$

The condition $\Delta u = 0$ must emerge from the bilinear form.

8.2.2 The even-order pattern

The same principle applies to a general elliptic equation of order $2r$. In multi-index notation, a divergence-form operator can be written schematically as

$$Lu = \sum_{|\alpha|,|\beta| \le r} (-1)^{|\alpha|} \partial^\alpha \left(a_{\alpha\beta}(x) \partial^\beta u \right), \tag{8.14}$$

with associated form

$$a(u, v) = \sum_{|\alpha|,|\beta| \le r} \int_\Omega a_{\alpha\beta}(x) \partial^\beta u \, \partial^\alpha v \, \mathrm{d}x. \tag{8.15}$$

Thus the energy space is contained in $H^r(\Omega)$. Boundary conditions involving derivatives of order at most $r - 1$ determine the closed subspace X; the remaining natural conditions are recovered from the weak form. This explains why an elliptic scalar equation treated by this energy method has even order.

8.3 The bilinear form for the biharmonic equation

Even when the differential equation is fixed, its bilinear form need not be unique. For the biharmonic equation, begin formally with $u, v \in C_c^\infty(\Omega)$. Integrating twice gives

$$a_H(u,v) = \int_\Omega \sum_{i,j=1}^{m} \partial_{ij} u \, \partial_{ij} v \, \mathrm{d}x = \int_\Omega (\Delta^2 u) v \, \mathrm{d}x. \tag{8.16}$$

Another possibility is

$$a_\Delta(u,v) = \int_\Omega \Delta u \, \Delta v \, \mathrm{d}x. \tag{8.17}$$

For compactly supported smooth functions, both forms represent the same interior operator. On spaces whose functions have nonzero boundary traces, their boundary terms differ. The boundary conditions therefore determine which form is appropriate.

For a_Δ, two integrations by parts give the exact identity

$$\begin{aligned} a_\Delta(u,v) &= -\int_\Omega \partial_i(\Delta u) \, \partial_i v \, \mathrm{d}x + \int_{\partial\Omega} (\partial_\nu v) \Delta u \, \mathrm{d}S \\ &= \int_\Omega (\Delta^2 u) v \, \mathrm{d}x - \int_{\partial\Omega} \partial_\nu(\Delta u) v \, \mathrm{d}S + \int_{\partial\Omega} (\partial_\nu v) \Delta u \, \mathrm{d}S. \end{aligned} \tag{8.18}$$

If $X = H_0^2(\Omega)$, then both v and $\partial_\nu v$ vanish on the boundary. The clamped conditions in (8.11) are therefore essential conditions.

If instead $X = H^2(\Omega) \cap H_0^1(\Omega)$, then $v = 0$ on the boundary but $\partial_\nu v$ remains free. The middle boundary term in (8.18) vanishes, and the last one forces

$$\Delta u = 0 \qquad \text{on } \partial\Omega.$$

Thus the second condition in (8.12) is a natural boundary condition, recovered after elliptic regularity makes the traces meaningful.

One can likewise build $\partial_\nu u = 0$ into the space while leaving the trace of u free. The corresponding free boundary trace in (8.18) then produces $\partial_\nu(\Delta u) = 0$ as the natural companion. In each case, boundary conditions and bilinear form must be chosen together.

8.4 Natural Neumann and Robin conditions

8.4.1 The Neumann condition

Return to the second-order problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ \partial_\nu u = 0 & \text{on } \partial\Omega. \end{cases} \tag{8.19}$$

The normal derivative cannot be encoded by the definition of an H^1 space, so take $X = H^1(\Omega)$ and

$$D(u,v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x. \tag{8.20}$$

For sufficiently regular u,

$$D(u,v) = \int_\Omega (-\Delta u)v \ \mathrm{d}x + \int_{\partial\Omega} (\partial_\nu u)v \ \mathrm{d}S. \tag{8.21}$$

The boundary term recovers $\partial_\nu u = 0$.

8.4.2 Producing a Robin condition

For

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ \partial_\nu u + u = 0 & \text{on } \partial\Omega, \end{cases} \tag{8.22}$$

the form (8.20) alone gives the wrong boundary term. The notes obtain the desired term by adding and subtracting the same first-order expression before integrating by parts. Formally, one writes

$$-\Delta u + a_i\partial_i u - a_i\partial_i u = f, \qquad -a_i\partial_i u = -\partial_i(a_i u) + (\partial_i a_i)u,$$

and chooses which copy of the first-order term to integrate by parts.

Choose a smooth vector field $a = (a_1, \ldots, a_m)$ satisfying

$$a \cdot \nu = 1 \qquad \text{on } \partial\Omega, \tag{8.23}$$

and define

$$B(u,v) = \int_\Omega \left[\nabla u \cdot \nabla v + (a_i\partial_i u)v + (a_i\partial_i v)u + (\partial_i a_i)uv\right] \ \mathrm{d}x. \tag{8.24}$$

The last three terms form a divergence:

$$(a_i\partial_i u)v + (a_i\partial_i v)u + (\partial_i a_i)uv = \partial_i(a_i uv).$$

Hence the divergence theorem and (8.23) give

$$\begin{aligned} B(u,v) &= \int_\Omega \nabla u \cdot \nabla v \ \mathrm{d}x + \int_{\partial\Omega} uv \ \mathrm{d}S \\ &= \int_\Omega (-\Delta u)v \ \mathrm{d}x + \int_{\partial\Omega} (\partial_\nu u + u)v \ \mathrm{d}S. \end{aligned} \tag{8.25}$$

Thus the Robin condition is the natural boundary condition for B.

8.5 The pure Neumann problem and the zero-mean space

The form (8.20) is not coercive on all of $H^1(\Omega)$: every constant function has zero energy. Indeed, the homogeneous Neumann problem has the constant solution $u = 1$. Assume for now that Ω is connected, and define the average

$$\langle w \rangle_\Omega = \frac{1}{|\Omega|} \int_\Omega w \ \mathrm{d}x$$

and the zero-mean space

$$\widetilde{H}^1(\Omega) = \left\{ w \in H^1(\Omega) : \langle w \rangle_\Omega = 0 \right\}. \tag{8.26}$$

Every $w \in H^1(\Omega)$ can be projected onto this space by

$$\widetilde{w} = w - \langle w \rangle_\Omega.$$

Poincaré's inequality on the zero-mean space states that

$$\|w\|_{L^2(\Omega)} \le C \|\nabla w\|_{L^2(\Omega)} \qquad (w \in \widetilde{H}^1(\Omega)).$$

Consequently,

$$D(w, w) = \int_\Omega |\nabla w|^2 \, \mathrm{d}x \ge c \|w\|^2_{H^1(\Omega)} \qquad (w \in \widetilde{H}^1(\Omega)), \tag{8.27}$$

so Lax–Milgram applies after the constants have been removed.

For arbitrary $f \in L^2(\Omega)$, let $u \in \widetilde{H}^1(\Omega)$ satisfy

$$\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x \qquad \text{for every } v \in \widetilde{H}^1(\Omega). \tag{8.28}$$

To recover the equation, take a smooth function φ and use

$$v = \varphi - \langle \varphi \rangle_\Omega \in \widetilde{H}^1(\Omega).$$

Since $\nabla v = \nabla \varphi$, equation (8.28) becomes

$$\begin{aligned} \int_\Omega \nabla u \cdot \nabla \varphi \, \mathrm{d}x &= \int_\Omega f \left(\varphi - \langle \varphi \rangle_\Omega \right) \mathrm{d}x \\ &= \int_\Omega \left(f - \langle f \rangle_\Omega \right) \varphi \, \mathrm{d}x. \end{aligned} \tag{8.29}$$

Thus the zero-mean weak formulation solves

$$-\Delta u = f - \langle f \rangle_\Omega, \qquad \partial_\nu u = 0. \tag{8.30}$$

The right-hand side is automatically orthogonal to the null space of the Laplacian.

It follows that the original Neumann problem (8.19) has a solution if and only if

$$\langle f \rangle_\Omega = 0. \tag{8.31}$$

When this compatibility condition holds, the solution is unique in $\widetilde{H}^1(\Omega)$. If one works in all of $H^1(\Omega)$, uniqueness is lost because an arbitrary constant may be added. On a disconnected domain, the corresponding compatibility and normalization are imposed separately on each connected component.

8.6 The Fredholm alternative

The Neumann problem illustrates a general linear-algebra principle. For a matrix $A : \mathbb{R}^n \to \mathbb{R}^n$,

$$\operatorname{Ran}(A) = \ker(A^*)^\perp. \tag{8.32}$$

Therefore $Ax = b$ is solvable precisely when

$$b \perp \ker(A^*).$$

This replaces invertibility when A has a nontrivial kernel.

To extend this statement to an infinite-dimensional Hilbert space, one seeks an operator that is an invertible operator plus a compact perturbation. After preconditioning by the invertible part, the equation takes the form

$$(I - T)u = f, \tag{8.33}$$

where T is compact. In elliptic problems, elliptic regularity provides a gain of derivatives, and compact Sobolev embedding then provides the required compactness. Hyperbolic equations do not generally have this elliptic gain of regularity, so this particular compactness argument is unavailable there.

Theorem 8.1 (Fredholm alternative)**.** *Let H be a Hilbert space and let $T : H \to H$ be compact. Then:*

(i) $\ker(I - T)$ *is finite-dimensional;*

(ii)

$$\dim \ker(I - T) = \dim \ker(I - T^*);$$

(iii)

$$\operatorname{Ran}(I - T) = \ker(I - T^*)^\perp. \tag{8.34}$$

In particular, $\operatorname{Ran}(I - T)$ *is closed.*

Consequently, (8.33) is solvable if and only if

$$f \perp \ker(I - T^*). \tag{8.35}$$

If the homogeneous equation $(I - T)u = 0$ has only the zero solution, then $I - T$ is invertible and (8.33) has a unique solution for every $f \in H$. Otherwise, solvability is governed by the orthogonality condition (8.35), exactly as in the finite-dimensional relation (8.32).

8.7 Further Elliptic Theory

The preceding chapters developed weak Dirichlet problems, Poincaré's inequality, Lax–Milgram existence, and elliptic regularity. We now record the parts of the supplementary elliptic theory that extend that account: negative-order data, shifted operators on the whole space, change of type, Gårding's inequality for lower-order terms, and a useful Hessian identity.

8.7.1 Negative-order data and shifted operators

The natural forcing space for a homogeneous Dirichlet problem is

$$H^{-1}(\Omega) := \left(H^1_0(\Omega)\right)^*.$$

Its norm is

$$\|F\|_{H^{-1}(\Omega)} := \sup_{\substack{\phi\in H^1_0(\Omega)\\ \phi\neq 0}} \frac{\left|\langle F,\phi\rangle\right|}{\left\|\phi\right\|_{H^1_0(\Omega)}}. \tag{8.36}$$

Every $f \in L^2(\Omega)$ defines an element of $H^{-1}(\Omega)$ through

$$\langle f,\phi\rangle = \int_\Omega f\phi \,\mathrm{d}x.$$

With the usual identification of $L^2(\Omega)$ with its dual, this gives the continuous Sobolev triple

$$H^1_0(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega). \tag{8.37}$$

The second inclusion is strict in general: not every bounded functional on $H^1_0(\Omega)$ is represented by an L^2 function.

The zero-order term in

$$-\Delta u + u = f \tag{8.38}$$

allows one to use the standard H^1 inner product

$$(u,v)_1 := \int_\Omega \left(\nabla u\cdot\nabla v + uv\right)\,\mathrm{d}x.$$

Unlike the unshifted Dirichlet problem, the Riesz argument for (8.38) does not require Poincaré's inequality. It therefore applies on the whole space. Since $H^1_0(\mathbb{R}^m) = H^1(\mathbb{R}^m)$, the map

$$-\Delta + I : H^1(\mathbb{R}^m) \longrightarrow H^{-1}(\mathbb{R}^m) \tag{8.39}$$

is onto and isometric when H^{-1} is equipped with the dual norm induced by $(\cdot,\cdot)_1$.

More generally, for $\mu > 0$ define

$$(u,v)_\mu := \int_\Omega \left(\nabla u\cdot\nabla v + \mu uv\right)\,\mathrm{d}x.$$

If $c = \max\{\mu,\mu^{-1}\}$, then

$$c^{-1}\,\|u\|_\mu^2 \le \|u\|_{H^1(\Omega)}^2 \le c\,\|u\|_\mu^2\,. \tag{8.40}$$

Thus every positive shift gives an equivalent Hilbert norm and, on a bounded domain with homogeneous Dirichlet data, a unique weak solution for every $f \in H^{-1}(\Omega)$.

8.7.2 An operator that changes type

Uniform ellipticity is a property of both the operator and the region on which it is studied. The Tricomi operator

$$L_T u = -\big(y\,\partial_x^2 u + \partial_y^2 u\big) \tag{8.41}$$

has principal form

$$y\xi_1^2 + \xi_2^2.$$

It is elliptic where $y > 0$, hyperbolic where $y < 0$, and degenerate on $y = 0$. On every strip

$$\big\{(x, y) : \varepsilon < y < 1\big\}\,, \qquad 0 < \varepsilon < 1,$$

it is uniformly elliptic. This example separates ordinary ellipticity from uniform ellipticity and shows how the same equation can change type across a hypersurface.

8.7.3 Lower-order terms and Gårding's inequality

Consider the divergence-form operator

$$Lu := -\partial_i\big(a_{ij}\partial_j u\big) + \partial_i(b_i u) + cu, \tag{8.42}$$

where $a_{ij} = a_{ji}$ and $a_{ij}, b_i, c \in L^\infty(\Omega)$. Assume there is $\theta > 0$ such that

$$a_{ij}(x)\xi_i\xi_j \geq \theta\,|\xi|^2 \quad \text{for almost every } x \in \Omega \text{ and every } \xi \in \mathbb{R}^m. \tag{8.43}$$

The weak bilinear form is

$$a(u, v) := \int_\Omega \big(a_{ij}\partial_j u\,\partial_i v - b_i u\,\partial_i v + cuv\big)\,\mathrm{d}x. \tag{8.44}$$

The term containing b_i generally makes this form nonsymmetric. The form is bounded on $H_0^1(\Omega) \times H_0^1(\Omega)$, but the lower-order terms may prevent direct coercivity.

Set

$$B := \sum_{i=1}^m \|b_i\|_{L^\infty(\Omega)}\,, \qquad C_0 := \|c\|_{L^\infty(\Omega)}\,.$$

For $u \in H_0^1(\Omega)$, uniform ellipticity gives

$$\theta\,\|\nabla u\|_{L^2}^2 \leq a(u, u) + B\,\|u\|_{L^2}\,\|\nabla u\|_{L^2} + C_0\,\|u\|_{L^2}^2\,.$$

Young's inequality yields, for every $\varepsilon > 0$,

$$\|u\|_{L^2}\,\|\nabla u\|_{L^2} \leq \varepsilon\,\|\nabla u\|_{L^2}^2 + \frac{1}{4\varepsilon}\,\|u\|_{L^2}^2\,.$$

Choosing ε so that $B\varepsilon < \theta/2$ and then using Poincaré's inequality gives constants $\beta > 0$ and $\gamma \geq 0$ such that

$$\beta\,\|u\|_{H_0^1(\Omega)}^2 \leq a(u, u) + \gamma\,\|u\|_{L^2(\Omega)}^2\,. \tag{8.45}$$

This is Gårding's inequality. It says precisely that coercivity can be restored by a sufficiently large positive shift.

Theorem 8.2 (Existence after a sufficiently large shift)**.** *Under the coefficient assumptions above, choose γ as in (8.45). For every $\mu \geq \gamma$ and every $f \in H^{-1}(\Omega)$, the problem*

$$Lu + \mu u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega$$

has a unique weak solution $u \in H_0^1(\Omega)$.

Proof. The shifted form

$$a_\mu(u, v) := a(u, v) + \mu(u, v)_{L^2(\Omega)}$$

is bounded, while (8.45) gives

$$\beta \|u\|_{H_0^1}^2 \leq a(u, u) + \gamma \|u\|_{L^2}^2 \leq a_\mu(u, u).$$

Lax–Milgram therefore gives the unique solution. □

8.7.4 A Hessian identity and the classical threshold

For $u \in C_c^\infty(\Omega)$, two integrations by parts give

$$\begin{aligned} \int_\Omega (\Delta u)^2 \, \mathrm{d}x &= \sum_{i,j=1}^m \int_\Omega \partial_i^2 u \, \partial_j^2 u \, \mathrm{d}x \\ &= \sum_{i,j=1}^m \int_\Omega \left|\partial_{ij} u\right|^2 \, \mathrm{d}x = \int_\Omega \left|D^2 u\right|^2 \, \mathrm{d}x. \end{aligned} \tag{8.46}$$

Consequently, if $-\Delta u = f$, then formally

$$\left\|D^2 u\right\|_{L^2(\Omega)} = \left\|f\right\|_{L^2(\Omega)}.$$

For an H^1 weak solution this calculation is not yet justified; the interior regularity theorem supplies the missing second derivatives.

More generally, if $a_{ij} \in C^{k+1}$, $f \in H^k$, and $u \in H^1$ is a weak solution of a uniformly elliptic principal-part equation, then

$$u \in H_{\mathrm{loc}}^{k+2}.$$

When $k > m/2$, Sobolev embedding gives $f \in C$ and $u \in C^2$, so the weak equation is a classical equation. Smooth coefficients and forcing then bootstrap the solution to C^∞ in the interior; the earlier boundary-regularity results give the corresponding conclusion up to a smooth boundary.

8.8 Problems

Problem 8.1. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary, let $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ be symmetric and positive definite, and let $c \in L^\infty(\Omega)$. Consider

$$Lu = -\operatorname{div}(A \nabla u) + c(x)u$$

and the bilinear form

$$B(u,v) := \int_\Omega (A\nabla u)\cdot\nabla v \;\mathrm{d}x + \int_\Omega c(x)uv \;\mathrm{d}x, \qquad u,v\in H_0^1(\Omega).$$

Let $\theta > 0$ be the smallest eigenvalue of A, and choose $C_P > 0$ so that

$$\|u\|^2_{L^2(\Omega)} \le C_P \, \|\nabla u\|^2_{L^2(\Omega)} \qquad (u \in H_0^1(\Omega)).$$

Assume that

$$c(x) \ge -\mu \quad \text{in } \Omega, \qquad 0 < \mu < \frac{\theta}{C_P}.$$

Show that B is bounded and coercive on $H_0^1(\Omega)$, and hence satisfies the hypotheses of the Lax–Milgram theorem.

Solution. First take smooth u and v. Integration by parts gives

$$\begin{aligned}\int_\Omega (Lu)v \;\mathrm{d}x &= -\int_\Omega \operatorname{div}\big(vA\nabla u\big)\;\mathrm{d}x + \int_\Omega \nabla v\cdot(A\nabla u)\;\mathrm{d}x + \int_\Omega c(x)uv\;\mathrm{d}x \\ &= -\int_{\partial\Omega} v(A\nabla u)\cdot\boldsymbol{n}\;\mathrm{d}S + B(u,v).\end{aligned}$$

Thus the boundary term vanishes when $v \in H_0^1(\Omega)$, which motivates the displayed bilinear form.

Because A is real and symmetric, it can be diagonalized as $A = V^\mathsf{T}\Sigma V$, where V is orthogonal and $\Sigma = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$. If $x = V\xi$, then

$$\xi^\mathsf{T}A\xi = x^\mathsf{T}\Sigma x = \sum_{i=1}^n \lambda_i x_i^2 \ge \lambda_{\min}\sum_{i=1}^n x_i^2 = \theta\,|\xi|^2.$$

Consequently,

$$\begin{aligned}\theta\,\|\nabla u\|^2_{L^2(\Omega)} &\le \int_\Omega \sum_{i,j=1}^n a_{ij}\,\partial_i u\,\partial_j u\;\mathrm{d}x \\ &= B(u,u) - \int_\Omega c(x)u^2\;\mathrm{d}x \\ &\le B(u,u) + \mu\int_\Omega u^2\;\mathrm{d}x \\ &\le B(u,u) + \mu C_P\,\|\nabla u\|^2_{L^2(\Omega)}.\end{aligned}$$

It follows that

$$B(u,u) \ge (\theta - \mu C_P)\,\|\nabla u\|^2_{L^2(\Omega)}.$$

Since $\theta - \mu C_P > 0$ and the gradient norm is equivalent to the $H_0^1(\Omega)$ norm, there is a $\beta > 0$ such that

$$B(u,u) \ge \beta\,\|u\|^2_{H_0^1(\Omega)}.$$

For boundedness, Cauchy–Schwarz and Poincaré's inequality yield

$$\begin{aligned}|B(u,v)| &\leq \sum_{i,j=1}^{n} \left|a_{ij}\right| \int_{\Omega} |\partial_i u| \left|\partial_j v\right| \, \mathrm{d}x + \|c\|_{L^\infty(\Omega)} \int_{\Omega} |u|\,|v| \, \mathrm{d}x \\ &\leq C \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}.\end{aligned}$$

Thus B is a bounded, coercive bilinear form on the Hilbert space $H_0^1(\Omega)$, so the Lax–Milgram theorem applies. □

Problem 8.2. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and $f \in L^2(\Omega)$. Prove that the Robin problem

$$\begin{cases}-\Delta u = f & \text{in } \Omega, \\ \partial_{\boldsymbol{n}} u + u = 0 & \text{on } \partial\Omega\end{cases}$$

has a unique weak solution $u \in H^1(\Omega)$.

Solution. For smooth u and v, the boundary condition gives

$$\begin{aligned}\int_{\Omega} (-\Delta u) v \, \mathrm{d}x &= \int_{\Omega} \nabla u \cdot \nabla v \, \mathrm{d}x - \int_{\partial\Omega} (\partial_{\boldsymbol{n}} u) v \, \mathrm{d}S \\ &= \int_{\Omega} \nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} uv \, \mathrm{d}S.\end{aligned}$$

Writing $T : H^1(\Omega) \to L^2(\partial\Omega)$ for the trace map, define

$$B(u,v) := \int_{\Omega} \nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} (Tu)(Tv) \, \mathrm{d}S.$$

The weak formulation is

$$B(u,v) = \int_{\Omega} fv \, \mathrm{d}x \qquad \text{for every } v \in H^1(\Omega).$$

The trace theorem and Cauchy–Schwarz imply

$$\begin{aligned}|B(u,v)| &\leq \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \|Tu\|_{L^2(\partial\Omega)} \|Tv\|_{L^2(\partial\Omega)} \\ &\leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)},\end{aligned}$$

so B is bounded.

To prove coercivity, suppose to the contrary that no positive coercivity constant exists. For every $k \in \mathbb{N}$ there is then a $u_k \in H^1(\Omega)$ such that

$$\|u_k\|_{H^1(\Omega)} = 1, \qquad B(u_k, u_k) < \frac{1}{k}.$$

After passing to a subsequence,

$$u_k \rightharpoonup u \quad \text{in } H^1(\Omega), \qquad u_k \to u \quad \text{in } L^2(\Omega),$$

where the second convergence follows from the Rellich compactness theorem. Moreover,

$$\int_\Omega |\nabla u_k|^2 \, dx \le B(u_k, u_k) \longrightarrow 0.$$

Thus $\nabla u = 0$ in the sense of distributions, so u is constant almost everywhere on each connected component of Ω. The strong L^2 convergence together with $\nabla u_k \to 0 = \nabla u$ in L^2 shows that $u_k \to u$ strongly in $H^1(\Omega)$. The trace theorem then gives $Tu_k \to Tu$ in $L^2(\partial\Omega)$, while

$$\|Tu_k\|^2_{L^2(\partial\Omega)} \le B(u_k, u_k) \longrightarrow 0.$$

Hence $Tu = 0$. Since u is constant on each component, this forces $u = 0$. On the other hand, strong convergence in $H^1(\Omega)$ gives

$$\|u\|_{H^1(\Omega)} = \lim_{k\to\infty} \|u_k\|_{H^1(\Omega)} = 1,$$

a contradiction. Therefore B is coercive on $H^1(\Omega)$.

Finally, the functional

$$v \longmapsto \int_\Omega f v \, dx$$

is bounded on $H^1(\Omega)$. Lax–Milgram therefore gives a unique weak solution $u \in H^1(\Omega)$. □

Problem 8.3. Let $\Omega \subset \mathbb{R}^n$ be bounded, let $1 < p < \infty$, and let $q = p/(p-1)$. Suppose $f \in L^q(\Omega)$ and $F \in C^1(\mathbb{R}^n)$ is convex with

$$|F(\xi)| \le C\big(1 + |\xi|^p\big), \qquad |DF(\xi)| \le C\big(1 + |\xi|^{p-1}\big).$$

On $W_0^{1,p}(\Omega)$ define

$$J(v) := \int_\Omega F(\nabla v) \, dx - \int_\Omega f v \, dx,$$

and assume that J is coercive. Prove that J is well defined and continuous, that it attains its infimum, and that every minimizer satisfies the weak Euler–Lagrange equation. Identify the associated Dirichlet problem.

Solution. For $v \in W_0^{1,p}(\Omega)$, the growth of F gives

$$\int_\Omega |F(\nabla v)| \, dx \le C\,|\Omega| + C \int_\Omega |\nabla v|^p \, dx < \infty.$$

Also, Hölder's inequality gives

$$\left| \int_\Omega f v \, dx \right| \le \|f\|_{L^q(\Omega)} \|v\|_{L^p(\Omega)} < \infty.$$

Thus J is well defined.

Suppose $v_m \to v$ in $W_0^{1,p}(\Omega)$. The mean value theorem and the growth of DF imply

$$|F(\nabla v_m) - F(\nabla v)| \le C\big(1 + |\nabla v_m|^{p-1} + |\nabla v|^{p-1}\big)\, |\nabla v_m - \nabla v|\,.$$

After integration and an application of Hölder's inequality,

$$\int_\Omega |F(\nabla v_m) - F(\nabla v)| \ \mathrm{d}x \le C\Big(|\Omega|^{1/q} + \|\nabla v_m\|_{L^p(\Omega)}^{p-1} + \|\nabla v\|_{L^p(\Omega)}^{p-1} \Big) \|\nabla v_m - \nabla v\|_{L^p(\Omega)} .$$

The right-hand side tends to zero. The linear term also converges because

$$\left| \int_\Omega f(v_m - v) \ \mathrm{d}x \right| \le \|f\|_{L^q(\Omega)} \|v_m - v\|_{L^p(\Omega)} \longrightarrow 0.$$

Hence $J(v_m) \to J(v)$.

Set

$$m := \inf_{v \in W_0^{1,p}(\Omega)} J(v)$$

and choose a minimizing sequence (v_k). Coercivity makes this sequence bounded in $W_0^{1,p}(\Omega)$. Since this is a reflexive Banach space, there is a subsequence and a $v^* \in W_0^{1,p}(\Omega)$ such that

$$v_k \rightharpoonup v^* \quad \text{in } W_0^{1,p}(\Omega).$$

The convex, continuous functional J is weakly lower semicontinuous, and therefore

$$m \le J(v^*) \le \liminf_{k\to\infty} J(v_k) = m.$$

Thus v^* is a minimizer. Convexity alone gives existence; if J is strictly convex, in particular if F is strictly convex, this minimizer is unique.

Let u be a minimizer and take $v \in W_0^{1,p}(\Omega)$. For every $t \in \mathbb{R}$,

$$J(u + tv) \ge J(u).$$

Moreover,

$$\frac{J(u + tv) - J(u)}{t} = \frac{1}{t} \int_\Omega \big[F(\nabla u + t\nabla v) - F(\nabla u)\big], \mathrm{d}x - \int_\Omega f v \ \mathrm{d}x.$$

For each x, the one-variable function $g_x(t) := F(\nabla u(x) + t\nabla v(x))$ satisfies

$$\frac{g_x(t) - g_x(0)}{t} = DF(\nabla u + c\nabla v) \cdot \nabla v$$

for some c between 0 and t. When $|t| \le 1$, the growth of DF and Young's inequality give an integrable bound, independent of t, of the form

$$\begin{aligned} |DF(\nabla u + c\nabla v) \cdot \nabla v| &\le C\big(1 + |\nabla u + c\nabla v|^{p-1}\big) |\nabla v| \\ &\le C'\big(1 + |\nabla u|^p + |\nabla v|^p\big). \end{aligned}$$

Dominated convergence therefore yields

$$\frac{\mathrm{d}}{\mathrm{d}t} J(u + tv)\Big|_{t=0} = \int_\Omega DF(\nabla u) \cdot \nabla v \ \mathrm{d}x - \int_\Omega f v \ \mathrm{d}x = 0.$$

Thus every minimizer satisfies

$$\int_\Omega DF(\nabla u)\cdot\nabla v\ \mathrm{d}x = \int_\Omega fv\ \mathrm{d}x \qquad \text{for every } v\in W_0^{1,p}(\Omega).$$

For smooth functions, integration by parts identifies the corresponding Dirichlet problem as

$$\begin{cases} -\operatorname{div}\big(DF(\nabla u)\big) = f & \text{in } \Omega,\\ u = 0 & \text{on } \partial\Omega.\end{cases}$$

□

Problem 8.4 (Analytic Cauchy data and the clamped biharmonic equation)**.** (a) Let $\mathcal{L}$ be a differential operator of order $2N$, let f be analytic, and let Ω be a bounded open set with analytic boundary. Consider the Cauchy data

$$\mathcal{L}u = f \quad \text{in } \Omega, \qquad u\big|_{\partial\Omega} = 0, \qquad \partial^\mu u\big|_{\partial\Omega} = 0 \quad (1\le|\mu|\le 2N-1).$$

Explain what the Cauchy–Kowalevsky theorem gives when its analytic and noncharacteristic hypotheses are satisfied.

(b) Assume that $\partial\Omega$ is smooth and that $f\in L^2(\Omega)$. Consider the clamped biharmonic problem

$$\Delta^2 u = f \quad \text{in } \Omega, \qquad u\in H_0^2(\Omega),$$

where

$$H_0^2(\Omega) = \left\{v\in H^2(\Omega) : v\big|_{\partial\Omega} = 0,\ \nabla v\big|_{\partial\Omega} = 0\right\}.$$

Give its weak formulation and establish existence, uniqueness, and the regularity recorded below.

Solution. For part (a), if the coefficients of $\mathcal{L}$, the forcing term, and the Cauchy hypersurface are analytic, and if $\partial\Omega$ is noncharacteristic for $\mathcal{L}$, the Cauchy–Kowalevsky theorem gives a unique analytic solution locally near each point of the hypersurface. The theorem is local; it does not, by itself, establish a single global solution throughout an arbitrary bounded domain.

For part (b), define on $H_0^2(\Omega)$ the bilinear form

$$D(u,v) := \int_\Omega \Delta u\,\Delta v\,\mathrm{d}x.$$

This form is coercive on $H_0^2(\Omega)$. Hence the Lax–Milgram theorem gives a unique $u\in H_0^2(\Omega)$ such that

$$D(u,v) = \int_\Omega fv\,\mathrm{d}x \qquad \text{for every } v\in H_0^2(\Omega).$$

Elliptic regularity then gives $u\in H^4(\Omega)$. For a smooth test function satisfying the clamped boundary conditions, two integrations by parts give

$$\begin{aligned} D(u,v) &= \int_\Omega (\Delta^2 u)v\,\mathrm{d}x - \int_{\partial\Omega}\partial_\nu(\Delta u)v\,\mathrm{d}S + \int_{\partial\Omega}(\Delta u)\partial_\nu v\,\mathrm{d}S\\ &= \int_\Omega fv\,\mathrm{d}x,\end{aligned}$$

because both boundary terms vanish. Thus the weak solution satisfies $\Delta^2 u = f$. □

Problem 8.5 (Gårding inequality for a uniformly elliptic operator)**.** Let

$$Lu = -\sum_{i,j} \partial_j\big(a_{ij}\partial_i u\big) + \sum_i b_i \partial_i u + cu,$$

where the coefficients are bounded and the principal coefficients satisfy the uniform ellipticity condition

$$\sum_{i,j} a_{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2, \qquad \theta > 0,$$

for almost every $x \in \Omega$ and every $\xi \in \mathbb{R}^n$. Define

$$D(u,v) := \sum_{i,j} \int_\Omega a_{ij}\partial_i u\, \partial_j v \,\mathrm{d}x + \sum_i \int_\Omega v b_i \partial_i u \,\mathrm{d}x + \int_\Omega cuv \,\mathrm{d}x.$$

Prove that there is $\gamma > 0$ such that

$$\|u\|^2_{H^1(\Omega)} \lesssim D(u,u) + \gamma \|u\|^2_{L^2(\Omega)}.$$

Solution. Set

$$\beta := \sum_i \|b_i\|_{L^\infty(\Omega)}, \qquad C := \operatorname*{ess\,inf}_{x\in\Omega} c(x).$$

Uniform ellipticity and Hölder's inequality give

$$\begin{aligned}
\theta \|\nabla u\|^2_{L^2(\Omega)} &\leq D(u,u) - \sum_i \int_\Omega b_i(\partial_i u)u \,\mathrm{d}x - \int_\Omega cu^2 \,\mathrm{d}x \\
&\leq D(u,u) + \beta \|\nabla u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} - C \|u\|^2_{L^2(\Omega)}.
\end{aligned}$$

For every $\varepsilon > 0$, Young's inequality yields

$$\beta \|\nabla u\|_{L^2} \|u\|_{L^2} \leq \beta\varepsilon \|\nabla u\|^2_{L^2} + \frac{\beta}{4\varepsilon} \|u\|^2_{L^2}.$$

Consequently,

$$\big(\theta - \beta\varepsilon\big) \|\nabla u\|^2_{L^2} + \left(C - \frac{\beta}{4\varepsilon}\right) \|u\|^2_{L^2} \leq D(u,u).$$

Choose $\varepsilon > 0$ so that $\theta - \beta\varepsilon > 0$, and then choose $\gamma > 0$ so that

$$\gamma + C - \frac{\beta}{4\varepsilon} > 0.$$

Adding $\gamma \|u\|^2_{L^2}$ to both sides gives

$$\|u\|^2_{H^1(\Omega)} \lesssim D(u,u) + \gamma \|u\|^2_{L^2(\Omega)}.$$

□

Chapter 9

Rellich Compactness, Spectral Theory, and Evolution Semigroups

9.1 From elliptic operators to evolution equations

Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with smooth boundary, so $\overline{\Omega}$ is compact. We now study the eigenvalue problem for the Laplacian and use its spectral resolution to solve the heat, Schrödinger, and wave equations.

The differential expression alone does not determine an operator. An operator consists of the expression together with its domain, and the domain records the boundary condition. For the Laplacian, the standard choices include Dirichlet, Neumann, and Robin boundary conditions. In particular, the Dirichlet realization of the positive Laplacian is

$$L = -\Delta, \qquad D(L) = H_0^1(\Omega) \cap H^2(\Omega). \tag{9.1}$$

On a smooth bounded domain, solving

$$-\Delta u = f, \qquad u|_{\partial\Omega} = 0,$$

gains regularity: for $f \in L^2(\Omega)$, the solution belongs to $H_0^1(\Omega) \cap H^2(\Omega)$. The embedding of this space into $L^2(\Omega)$ is compact. Thus, whenever zero is not an eigenvalue, L^{-1} is compact as an operator on $L^2(\Omega)$. The compactness behind this observation is Rellich compactness.

9.2 Rellich compactness

For $1 \leq p < m$, define the critical Sobolev exponent

$$p^* := \frac{mp}{m-p}.$$

We write $X \Subset Y$ when the embedding $X \hookrightarrow Y$ is compact.

Theorem 9.1 (Rellich compactness)**.** *Let $\Omega \subset \mathbb{R}^m$ be bounded with smooth boundary, and let $1 \leq p < m$. Then*

$$W^{1,p}(\Omega) \Subset L^q(\Omega) \qquad \text{for every } 1 \leq q < p^*. \tag{9.2}$$

The endpoint $p = m$ follows from the subcritical theorem. Since Ω is bounded, Hölder's inequality gives

$$W^{1,m}(\Omega) \hookrightarrow W^{1,p}(\Omega) \qquad (1 \le p < m).$$

Moreover, $p^* \to \infty$ as $p \uparrow m$. Given any finite $q \ge 1$, choose $p < m$ sufficiently close to m that $q < p^*$. Then

$$W^{1,m}(\Omega) \hookrightarrow W^{1,p}(\Omega) \Subset L^q(\Omega).$$

Consequently,

$$W^{1,m}(\Omega) \Subset L^q(\Omega) \qquad \text{for every } 1 \le q < \infty. \tag{9.3}$$

In particular, one obtains compactness in the same exponent for every finite p:

$$W^{1,p}(\Omega) \Subset L^p(\Omega), \qquad 1 \le p < \infty. \tag{9.4}$$

Indeed, if $p < m$, then $p < p^*$ and this is (9.2); if $p = m$, it is (9.3). If $p > m$, choose $r < m$ so close to m that $p < r^*$. Boundedness of Ω gives

$$W^{1,p}(\Omega) \hookrightarrow W^{1,r}(\Omega) \Subset L^p(\Omega).$$

The critical exponent is not compact

For $1 \le p < m$, the Sobolev embedding

$$W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$$

is continuous, but in general it is not compact. Compactness holds only below the critical exponent.

9.3 Self-adjoint elliptic operators

Let A be a densely defined operator on a Hilbert space. If A is self-adjoint, then $A = A^*$ and, in particular,

$$\langle Au, v\rangle = \langle u, Av\rangle \qquad (u, v \in D(A)).$$

For the Dirichlet Laplacian in (9.1), integration by parts gives

$$\begin{aligned}\langle -\Delta u, v\rangle_{L^2} &= \int_\Omega \nabla u \cdot \overline{\nabla v}\, \mathrm{d}x \\ &= \langle u, -\Delta v\rangle_{L^2}, \qquad u, v \in H_0^1(\Omega) \cap H^2(\Omega).\end{aligned}$$

Thus the Dirichlet realization is self-adjoint. Its eigenvalues are real, and eigenfunctions belonging to distinct eigenvalues are orthogonal in $L^2(\Omega)$.

9.3.1 The first Dirichlet eigenvalue

The first eigenfunction is obtained by minimizing the Dirichlet energy under an L^2 constraint. Set

$$\mu_1 := \inf_{\substack{u\in H_0^1(\Omega)\\ \|u\|_{L^2}=1}} \frac{1}{2}\int_\Omega |\nabla u|^2 \, \mathrm{d}x. \tag{9.5}$$

Let (u_j) be a minimizing sequence. The constraint and the bounded energy imply that (u_j) is bounded in $H_0^1(\Omega)$. After taking a subsequence,

$$u_j \rightharpoonup e_1 \quad \text{weakly in } H_0^1(\Omega).$$

Rellich compactness gives

$$u_j \longrightarrow e_1 \quad \text{strongly in } L^2(\Omega).$$

Hence $\|e_1\|_{L^2} = 1$. By weak lower semicontinuity,

$$\frac{1}{2}\int_\Omega |\nabla e_1|^2 \, \mathrm{d}x \le \liminf_{j\to\infty} \frac{1}{2}\int_\Omega |\nabla u_j|^2 \, \mathrm{d}x = \mu_1.$$

Since e_1 is admissible, equality holds. Thus the infimum is attained. Equality of the gradient norms, together with weak convergence, also gives strong convergence in $H_0^1(\Omega)$.

The Lagrange-multiplier equation for this constrained minimum is

$$\int_\Omega \nabla e_1 \cdot \nabla\varphi \, \mathrm{d}x = \lambda_1 \int_\Omega e_1\varphi \, \mathrm{d}x, \qquad \varphi \in H_0^1(\Omega),$$

where

$$\lambda_1 = \int_\Omega |\nabla e_1|^2 \, \mathrm{d}x = 2\mu_1. \tag{9.6}$$

Equivalently,

$$-\Delta e_1 = \lambda_1 e_1$$

in the weak sense. Elliptic regularity then places e_1 in the operator domain.

Poincaré's inequality gives $\lambda_1 > 0$. More generally, define successively

$$\lambda_k := \inf_{\substack{u\in H_0^1(\Omega),\ \|u\|_{L^2}=1\\ \langle u,e_j\rangle_{L^2}=0,\ 1\le j<k}} \int_\Omega |\nabla u|^2 \, \mathrm{d}x. \tag{9.7}$$

The same compactness argument produces an eigenfunction e_k. Consequently,

$$0 < \lambda_1 \le \lambda_2 \le \cdots, \qquad \lambda_k \longrightarrow \infty,$$

and the e_k may be chosen to form an orthonormal basis of $L^2(\Omega)$. This variational construction extends to self-adjoint elliptic operators whose quadratic forms are bounded from below; ellipticity supplies the required lower bound, possibly after shifting the operator by a constant.

9.3.2 Neumann boundary conditions

For the Neumann Laplacian, minimize the same energy over $H^1(\Omega)$ with $\|u\|_{L^2} = 1$. Constant functions have zero energy, so the first eigenvalue is

$$\lambda_1^N = 0, \qquad e_1^N = |\Omega|^{-1/2}.$$

If Ω is connected, the kernel consists exactly of the constant functions. The next eigenvalue is obtained by restricting to the mean-zero subspace,

$$\int_\Omega u \ \mathrm{d}x = 0,$$

and the mean-zero Poincaré inequality gives $\lambda_2^N > 0$.

9.4 Abstract evolution and the resolvent

Consider the abstract initial-value problem

$$\partial_t u = Au, \qquad u(0) = u_0, \tag{9.8}$$

where A is a matrix or, more generally, an operator. Formally,

$$u(t) = \mathrm{e}^{tA} u_0.$$

If $\mathcal{L}$ denotes the Laplace transform in time, then

$$s\mathcal{L}u - u_0 = A\mathcal{L}u,$$

and therefore

$$\mathcal{L}u(s) = (sI - A)^{-1} u_0. \tag{9.9}$$

Inverting the transform along a vertical line to the right of the spectrum gives the formal resolvent representation

$$u(t) = \left(\frac{1}{2\pi \mathrm{i}} \int_{\gamma - \mathrm{i}\infty}^{\gamma + \mathrm{i}\infty} \mathrm{e}^{st} (sI - A)^{-1} \ \mathrm{d}s \right) u_0. \tag{9.10}$$

The principal examples are

$$\partial_t u = \Delta u, \qquad \partial_t u = \mathrm{i}\Delta u, \qquad \partial_t^2 u - \Delta u = 0,$$

namely the heat, Schrödinger, and wave equations. On a bounded domain, boundary conditions must be included in the domain of the operator.

9.5 Bounded domains versus the whole space

Compactness of the domain is decisive for the spectral picture. On a bounded interval, the Dirichlet Laplacian has a discrete sequence of eigenvalues. On the whole line, consider

$$-\partial_x^2 u = \lambda u, \qquad u \in L^2(\mathbb{R}).$$

The three cases are:

$$\begin{aligned}
\lambda = 0: &\quad u(x) = a + bx,\\
\lambda > 0: &\quad u(x) = A\cos(\sqrt{\lambda}\,x) + B\frac{\sin(\sqrt{\lambda}\,x)}{\sqrt{\lambda}},\\
\lambda < 0: &\quad u(x) = A\cosh(\sqrt{-\lambda}\,x) + B\frac{\sinh(\sqrt{-\lambda}\,x)}{\sqrt{-\lambda}}.
\end{aligned}$$

None of these is in $L^2(\mathbb{R})$ unless it is identically zero. Thus $-\partial_x^2$ has no L^2 eigenfunctions on the whole line. Nevertheless,

$$\sigma(-\partial_x^2) = [0, \infty).$$

For every $\lambda \geq 0$, one can construct normalized approximate eigenfunctions for which

$$\|(-\partial_x^2 - \lambda)u_j\|_{L^2} \longrightarrow 0.$$

This is continuous spectrum rather than point spectrum. In particular, zero lies in the spectrum, so the inverse of $-\partial_x^2$ is not a bounded operator on $L^2(\mathbb{R})$; no compact inverse or discrete eigenfunction expansion is available.

9.6 The Dirichlet heat equation on an interval

Consider

$$\begin{cases} \partial_t u = \partial_x^2 u, & 0 < x < 1,\ t > 0,\\ u(0,t) = u(1,t) = 0, & t > 0,\\ u(x,0) = u_0(x), & 0 < x < 1. \end{cases} \tag{9.11}$$

The normalized Dirichlet eigenfunctions and eigenvalues are

$$e_n(x) = \sqrt{2}\sin(n\pi x), \qquad -\partial_x^2 e_n = n^2\pi^2 e_n, \qquad n = 1, 2, \ldots.$$

The functions $(e_n)_{n\geq 1}$ form an orthonormal basis of $L^2(0,1)$. If $u_0 \in L^2(0,1)$, write

$$u_0 = \sum_{n=1}^{\infty} a_n e_n, \qquad a_n = \int_0^1 u_0(x)e_n(x)\,\mathrm{d}x, \qquad \sum_{n=1}^{\infty} |a_n|^2 < \infty.$$

Separation of variables gives

$$u(x,t) = \sum_{n=1}^{\infty} a_n \mathrm{e}^{-n^2\pi^2 t} e_n(x), \qquad t \geq 0. \tag{9.12}$$

An arbitrary L^2 initial value need not have a boundary trace, so the boundary condition at $t = 0$ need not be meaningful. By contrast,

$$u_0 \in H_0^1(0,1) \qquad \Longleftrightarrow \qquad \sum_{n=1}^{\infty} n^2\pi^2|a_n|^2 < \infty.$$

Even when u_0 belongs only to L^2, the solution belongs to $H_0^1(0,1)$ for every $t > 0$.

The initial condition is attained strongly in L^2:

$$\|u(\cdot,t) - u_0\|_{L^2}^2 = \sum_{n=1}^{\infty} |\mathrm{e}^{-n^2\pi^2 t} - 1|^2|a_n|^2 \longrightarrow 0 \qquad (t \downarrow 0).$$

Thus

$$u \in C([0,\infty); L^2(0,1)) \cap C((0,\infty); H_0^1(0,1)).$$

9.6.1 Instantaneous smoothing

For $t > 0$, the series may be differentiated term by term:

$$\partial_x u(x,t) = \sqrt{2}\sum_{n=1}^{\infty} n\pi a_n \mathrm{e}^{-n^2\pi^2 t}\cos(n\pi x),$$
$$\partial_x^2 u(x,t) = -\sqrt{2}\sum_{n=1}^{\infty} n^2\pi^2 a_n \mathrm{e}^{-n^2\pi^2 t}\sin(n\pi x).$$

The exponential factor dominates every power of n. Repeating the argument gives derivatives of every order and hence

$$u \in C^\infty((0,\infty) \times [0,1]).$$

This is the regularizing effect of the heat equation.

9.7 The heat semigroup

Define

$$S(t)u_0 := \sum_{n=1}^{\infty} a_n \mathrm{e}^{-n^2\pi^2 t} e_n, \qquad t \geq 0. \tag{9.13}$$

Then

$$S(0) = I, \qquad S(t_1)S(t_2) = S(t_1 + t_2) \quad (t_1, t_2 \geq 0).$$

Moreover,

$$\|S(t)u_0 - u_0\|_{L^2} \longrightarrow 0 \qquad (t \downarrow 0),$$

so $(S(t))_{t\geq 0}$ is a C_0-semigroup. Strong convergence implies weak convergence:

$$\langle S(t)u_0, v\rangle_{L^2} \longrightarrow \langle u_0, v\rangle_{L^2} \qquad (v \in L^2(0,1)).$$

However, convergence does not hold in operator norm. Indeed, for every $t > 0$,

$$\|S(t) - I\|_{L^2 \to L^2} = \sup_{n \geq 1} |\mathrm{e}^{-n^2\pi^2 t} - 1| = 1.$$

Let $A = \partial_x^2$ with

$$D(A) = H_0^1(0,1) \cap H^2(0,1).$$

For every $u_0 \in L^2(0,1)$ and $t > 0$, smoothing gives

$$\frac{\mathrm{d}}{\mathrm{d}t} S(t)u_0 = AS(t)u_0.$$

The derivative at $t = 0$ exists in L^2 precisely when $u_0 \in D(A)$, and then

$$\lim_{t \downarrow 0} \frac{S(t)u_0 - u_0}{t} = Au_0.$$

Thus the generator A records exactly the initial data for which the evolution equation holds classically at the initial time.

9.8 The Schrödinger group and loss of smoothing

On the same interval with Dirichlet boundary conditions, consider

$$\partial_t u = \mathrm{i}\partial_x^2 u, \qquad u(0) = u_0 = \sum_{n=1}^{\infty} a_n e_n.$$

The solution is

$$U(t)u_0 = \sum_{n=1}^{\infty} a_n \mathrm{e}^{-\mathrm{i}n^2\pi^2 t} e_n, \qquad t \in \mathbb{R}. \tag{9.14}$$

Since every multiplier has modulus one, $U(t)$ is defined for all positive and negative times, and

$$U(t_1)U(t_2) = U(t_1 + t_2), \qquad \|U(t)u_0\|_{L^2} = \|u_0\|_{L^2}.$$

Hence $(U(t))_{t \in \mathbb{R}}$ is a unitary C_0-group.

Unlike the heat multipliers, the Schrödinger multipliers do not decay as $n \to \infty$. Therefore the flow gives no gain of regularity. If u_0 belongs only to L^2, its boundary trace need not be defined, and the evolution does not repair this. If $u_0 \in H_0^1(0,1)$, then the Dirichlet trace is meaningful and is preserved by the flow.

Remark 9.2. The heat fundamental solution is positive, and heat evolution preserves positivity. This positivity is closely related to the maximum principle.

Chapter 10

Fourier Methods for Evolution Equations

10.1 Evolution equations on Euclidean space

Evolution equations on all of $\mathbb{R}^d$ are especially amenable to Fourier analysis. The Fourier transform turns constant-coefficient spatial differential operators into multiplication operators, so that many partial differential equations become ordinary differential equations in time. The discussion here follows the Fourier-space viewpoint used in Rauch's treatment.

We begin by recalling the function spaces in which this method is carried out. The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is the space of smooth functions for which every seminorm

$$N_p(\varphi) := \sup_{\substack{x\in\mathbb{R}^d \\ |\alpha|+|\beta|\le p}} \left|x^\alpha \partial^\beta \varphi(x)\right| \tag{10.1}$$

is finite. Its continuous dual $\mathcal{S}'(\mathbb{R}^d)$ is the space of tempered distributions, and the Fourier transform on $\mathcal{S}'$ is defined by duality. Since $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$, every tempered distribution determines a distribution:

$$\mathcal{S}'(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d).$$

Tempered distributions are distinguished by their controlled behavior at spatial infinity.

For $s \in \mathbb{R}$, define

$$H^s(\mathbb{R}^d) := \left\{ u \in \mathcal{S}'(\mathbb{R}^d) : \int_{\mathbb{R}^d} (1+|\xi|^2)^s |\widehat{u}(\xi)|^2 \, \mathrm{d}\xi < \infty \right\}. \tag{10.2}$$

For $s \ge 0$, elements of H^s are L^2 functions. When $s < 0$, they may be genuine tempered distributions.

We retain the Fourier convention introduced earlier in the book:

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi \mathrm{i} x\cdot\xi} \, \mathrm{d}x, \qquad f(x) = \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi \mathrm{i} x\cdot\xi} \, \mathrm{d}\xi. \tag{10.3}$$

Consequently,

$$\widehat{\partial_j f}(\xi) = 2\pi \mathrm{i} \xi_j \widehat{f}(\xi), \qquad \widehat{\Delta f}(\xi) = -4\pi^2 |\xi|^2 \widehat{f}(\xi). \tag{10.4}$$

10.2 The Picard-iteration model

Before turning to partial differential equations, consider the autonomous ordinary differential equation

$$\dot{x} = F(x), \qquad x(0) = x_0, \qquad x(t) \in \mathbb{R}^d, \tag{10.5}$$

where F is locally Lipschitz. Picard iteration starts with $x^{(0)}(t) \equiv x_0$ and defines

$$x^{(m+1)}(t) = x_0 + \int_0^t F\big(x^{(m)}(\sigma)\big)\, \mathrm{d}\sigma. \tag{10.6}$$

The natural space for convergence is

$$C\big([0,T];\mathbb{R}^d\big),$$

rather than an L^1 or L^2 space: the equation evaluates the unknown pointwise in time. Choose a ball $B_r(x_0)$ on which F has Lipschitz constant L and bound $|F| \le M$. If $MT \le r$, the Picard map sends paths taking values in this ball back into the same ball. Set

$$d_m(T) := \sup_{0\le t\le T} \left|x^{(m+1)}(t) - x^{(m)}(t)\right|. \tag{10.7}$$

Then

$$\begin{aligned}\left|x^{(m+1)}(t) - x^{(m)}(t)\right| &\le \int_0^t \left|F\big(x^{(m)}(\sigma)\big) - F\big(x^{(m-1)}(\sigma)\big)\right| \,\mathrm{d}\sigma \\ &\le L\int_0^t \left|x^{(m)}(\sigma) - x^{(m-1)}(\sigma)\right| \,\mathrm{d}\sigma,\end{aligned}$$

and hence

$$d_m(T) \le LT\, d_{m-1}(T). \tag{10.8}$$

For example, if $LT \le \frac{1}{4}$, then

$$d_m(T) \le 4^{-m} d_0(T).$$

The geometric-series estimate

$$\sup_{0\le t\le T} |x^{(m)}(t) - x_0| \le \sum_{j=0}^{m-1} d_j(T) \le \frac{4}{3} d_0(T)$$

both keeps the iterates in the chosen ball, after T is reduced if necessary, and proves uniform convergence.

For an evolution equation, the analogous path space is

$$X_T^s := C\big([0,T];H^s(\mathbb{R}^d)\big), \qquad \|u\|_{X_T^s} := \sup_{0\le t\le T} \|u(t)\|_{H^s}. \tag{10.9}$$

Because H^s is a Banach space, so is X_T^s. If $u \in X_T^s$, the Banach-valued Riemann integral

$$\int_a^b u(\sigma)\,\mathrm{d}\sigma \in H^s, \qquad a,b \in [0,T],$$

is well defined. This is the setting in which a Duhamel or Picard formula can be interpreted.

10.3 The heat equation

Consider

$$\partial_t u = \Delta u, \qquad x \in \mathbb{R}^d, \qquad u(0) = u_0. \tag{10.10}$$

A direct Picard scheme of the form

$$\partial_t u^{(m+1)} = \Delta u^{(m)}$$

does not close in H^s: if $u^{(m)} \in H^s$, then generally $\Delta u^{(m)} \in H^{s-2}$. Two derivatives are lost at each step.

Fourier transformation avoids this loss. If $u \in C([0,T];H^s)$ solves (10.10), then $\partial_t u = \Delta u \in C([0,T];H^{s-2})$, so $u \in C^1([0,T];H^{s-2})$ and the time derivative may be transformed. Equations (10.10) and (10.4) give

$$\partial_t \widehat{u}(\xi,t) = -4\pi^2|\xi|^2\widehat{u}(\xi,t). \tag{10.11}$$

Thus

$$\boxed{\widehat{u}(\xi,t) = e^{-4\pi^2 t|\xi|^2}\widehat{u}_0(\xi)}. \tag{10.12}$$

For $t \geq 0$, the multiplier has modulus at most one. Hence every $u_0 \in H^s$, in particular every $u_0 \in H^s$ with $s \geq 0$, gives a solution

$$u \in C\big([0,\infty);H^s(\mathbb{R}^d)\big).$$

For every $t > 0$, the Gaussian decay in frequency gives $u(t) \in H^r$ for every r, so u is smooth in x; the equation then gives smoothness in t as well.

For $t < 0$, the factor in (10.12) grows exponentially in $|\xi|^2$. It therefore does not define a bounded solution operator on H^s, and backward evolution is not available for arbitrary H^s data. This is the basic backward ill-posedness of the heat equation.

Taking the inverse Fourier transform yields the heat kernel

$$K_t(x) = \frac{1}{(4\pi t)^{d/2}}e^{-|x|^2/(4t)}, \qquad t > 0, \tag{10.13}$$

and the convolution formula

$$u(x,t) = (K_t * u_0)(x) = \int_{\mathbb{R}^d} K_t(x-y)u_0(y)\,\mathrm{d}y. \tag{10.14}$$

10.4 The Schrödinger equation

Consider the free Schrödinger equation

$$\mathrm{i}\partial_t u - \Delta u = 0, \qquad u(0) = u_0. \tag{10.15}$$

Its Fourier transform satisfies

$$\mathrm{i}\partial_t\widehat{u} + 4\pi^2|\xi|^2\widehat{u} = 0,$$

or equivalently

$$\partial_t \widehat{u} - 4\pi^2 \mathrm{i}|\xi|^2 \widehat{u} = 0.$$

Therefore

$$\boxed{\widehat{u}(\xi, t) = e^{4\pi^2 \mathrm{i}t|\xi|^2} \widehat{u}_0(\xi)}. \tag{10.16}$$

The multiplier has modulus one, so

$$\|u(t)\|_{H^s} = \|u_0\|_{H^s}, \qquad t \in \mathbb{R}. \tag{10.17}$$

Thus $u_0 \in H^s$ gives $u \in C(\mathbb{R}; H^s)$, but the solution generally has no more Sobolev regularity than the initial datum.

For Schwartz initial data and $t \neq 0$, the solution can also be written as an oscillatory convolution

$$u(x, t) = (S_t * u_0)(x), \qquad S_t(x) = (-4\pi \mathrm{i}t)^{-d/2} e^{-\mathrm{i}|x|^2/(4t)}, \tag{10.18}$$

where the complex power is taken with the oscillatory branch for which $\widehat{S_t}(\xi) = e^{4\pi^2 \mathrm{i}t|\xi|^2}$.

Remark 10.1 (Uniqueness). Uniqueness for both the heat and Schrödinger equations is immediate in the spaces above. If two solutions have the same initial datum, their difference has zero initial datum. After Fourier transformation, each frequency solves a scalar ordinary differential equation with zero initial value and is therefore identically zero.

10.5 The wave equation

The Cauchy problem for the wave equation is

$$\partial_t^2 u - \Delta u = 0, \qquad u(0) = u_0, \qquad \partial_t u(0) = u_1. \tag{10.19}$$

Fourier transformation in x gives

$$\partial_t^2 \widehat{u} + 4\pi^2 |\xi|^2 \widehat{u} = 0.$$

Solving this ordinary differential equation yields

$$\boxed{\widehat{u}(\xi, t) = \cos(2\pi|\xi|t)\widehat{u}_0(\xi) + \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\widehat{u}_1(\xi)}. \tag{10.20}$$

The quotient is defined continuously at $\xi = 0$, where its value is t.

The first multiplier in (10.20) is bounded, so $u_0 \in H^s$ gives a contribution in H^s. The second multiplier gains one derivative at high frequency without acquiring a singularity at the origin. Indeed,

$$\left|\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\right| \leq \min\left\{|t|, \frac{1}{2\pi|\xi|}\right\}.$$

This low-frequency behavior matters. For example, if $f \in H^s$, then the function with Fourier transform

$$\frac{1}{1 + |\xi|}\widehat{f}(\xi)$$

belongs to H^{s+1}, whereas the function suggested by $|\xi|^{-1}\widehat{f}(\xi)$ need not, because of the singularity at $\xi = 0$. The sine multiplier avoids this problem. Consequently, for $s \geq 1$,

$$u_0 \in H^s, \qquad u_1 \in H^{s-1} \quad \Longrightarrow \quad u \in C(\mathbb{R}; H^s) \cap C^1(\mathbb{R}; H^{s-1}). \tag{10.21}$$

This is the PDE analogue of replacing a second-order ordinary differential equation by a first-order system. For example,

$$\ddot{x} + a\dot{x} + cx = 0, \qquad y = \dot{x},$$

becomes

$$\dot{x} = y, \qquad \dot{y} = -ay - cx.$$

For the wave equation, the state is

$$W(t) = \begin{pmatrix} u(t) \\ \partial_t u(t) \end{pmatrix} \in H^s \times H^{s-1}. \tag{10.22}$$

10.5.1 Physical-space wave kernels

Let R_t be the distribution whose Fourier transform is

$$\widehat{R_t}(\xi) = \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}. \tag{10.23}$$

For $t > 0$, its action on a suitable function g is

$$(R_t * g)(x) = \begin{cases} \dfrac{1}{2}\displaystyle\int_{x-t}^{x+t} g(y)\,\mathrm{d}y, & d = 1, \\ \dfrac{1}{2\pi}\displaystyle\int_{|x-y|<t} \frac{g(y)}{\sqrt{t^2 - |x-y|^2}}\,\mathrm{d}y, & d = 2, \\ \dfrac{1}{4\pi t}\displaystyle\int_{|x-y|=t} g(y)\,\mathrm{d}S(y), & d = 3. \end{cases} \tag{10.24}$$

Since

$$\partial_t \widehat{R_t}(\xi) = \cos(2\pi|\xi|t),$$

the full solution is

$$u(t) = \partial_t R_t * u_0 + R_t * u_1. \tag{10.25}$$

The three model equations

With the Fourier convention (10.3), the three solution multipliers are

$$\begin{aligned} \partial_t u - \Delta u = 0 &: \quad e^{-4\pi^2 t|\xi|^2}, \\ \mathrm{i}\partial_t u - \Delta u = 0 &: \quad e^{4\pi^2 \mathrm{i} t|\xi|^2}, \\ \partial_t^2 u - \Delta u = 0 &: \quad \cos(2\pi|\xi|t), \qquad \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}. \end{aligned}$$

The heat flow smooths and evolves only forward on arbitrary Sobolev data. The Schrödinger flow preserves Sobolev regularity for all $t \in \mathbb{R}$. The wave flow uses the natural data space $H^s \times H^{s-1}$ and propagates one position datum together with one velocity datum.

10.6 A heat equation with a potential

Consider now

$$\partial_t u = \Delta u + V(x)u, \qquad u(0) = u_0. \tag{10.26}$$

The Fourier transform no longer diagonalizes the whole right-hand side, because multiplication by V in physical space becomes convolution in frequency. There are two natural approaches.

10.6.1 The resolvent approach

Set $A = \Delta + V(x)$ and study its spectrum. Formally, if

$$\widetilde{u}(\lambda) = \int_0^\infty e^{-\lambda t} u(t)\, \mathrm{d}t$$

is the Laplace transform, then

$$(\lambda - \Delta - V)\widetilde{u}(\lambda) = u_0, \qquad \widetilde{u}(\lambda) = (\lambda - \Delta - V)^{-1} u_0. \tag{10.27}$$

Thus the evolution can be studied through the resolvent and spectral properties of A.

10.6.2 Variation of parameters and Duhamel's formula

For the ordinary differential equation

$$\dot{x} = Ax + f(t, x(t)), \qquad x(0) = x_0, \tag{10.28}$$

variation of parameters gives

$$x(t) = e^{tA} x_0 + \int_0^t e^{(t-\sigma)A} f(\sigma, x(\sigma))\, \mathrm{d}\sigma. \tag{10.29}$$

One may therefore iterate the integral equation:

$$x^{(m+1)}(t) = e^{tA} x_0 + \int_0^t e^{(t-\sigma)A} f\big(\sigma, x^{(m)}(\sigma)\big)\, \mathrm{d}\sigma. \tag{10.30}$$

This formulation is useful precisely when the propagator e^{tA} does not lose regularity.

For (10.26), the naive differential iteration

$$\partial_t u^{(m+1)} = \Delta u^{(m)} + V u^{(m)}$$

again loses two derivatives. Duhamel's formula instead suggests

$$u^{(m+1)}(t) = e^{t\Delta}u_0 + \int_0^t e^{(t-\sigma)\Delta} V u^{(m)}(\sigma)\, \mathrm{d}\sigma. \tag{10.31}$$

Here $e^{t\Delta}$ is the heat propagator, with Fourier multiplier $e^{-4\pi^2 t|\xi|^2}$, and

$$\|e^{t\Delta} f\|_{H^s} \le \|f\|_{H^s}, \qquad t \ge 0. \tag{10.32}$$

To close the iteration in H^s, assume that multiplication by V is bounded on H^s; write its operator norm as M_V. For example, when s is a nonnegative integer, sufficiently many bounded derivatives of V provide this property. If Φ denotes the right-hand side of (10.31), then

$$\|\Phi u - \Phi v\|_{X_T^s} \le M_V T\, \|u - v\|_{X_T^s}. \tag{10.33}$$

For $M_V T < 1$, this is a contraction on X_T^s, exactly as in the ODE argument.

The same formula gives

$$\|u(t)\|_{H^s} \le \|u_0\|_{H^s} + M_V \int_0^t \|u(\sigma)\|_{H^s}\, \mathrm{d}\sigma, \tag{10.34}$$

and hence

$$\|u(t)\|_{H^s} \le e^{M_V t}\|u_0\|_{H^s}.$$

This bound supports the continuation argument in the notes. If a maximal solution had a finite endpoint T, the bound and the Duhamel formula would give a limit $u(T)$ in H^s as $t \uparrow T$. One could then restart from $u(T)$ using

$$u(t) = e^{(t-T)\Delta}u(T) + \int_T^t e^{(t-\sigma)\Delta} V u(\sigma)\, \mathrm{d}\sigma. \tag{10.35}$$

The same local contraction applies beyond T, contradicting maximality. Under the stated multiplier assumption on V, the solution therefore continues for every $t \ge 0$.

10.7 Problems

Problem 10.1 (A variational solution of a semilinear Dirichlet problem)**.** Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with smooth boundary, and set

$$\lambda := \inf\left\{ \int_\Omega |\nabla u|^2\, \mathrm{d}x : u \in H_0^1(\Omega), \quad \|u\|_{L^4(\Omega)} = 1 \right\}.$$

(a) Show that the infimum is attained by some $u^* \in H_0^1(\Omega)$.
(b) Derive the Euler–Lagrange equation for u^* and rescale it to obtain a weak solution of

$$\begin{cases} -\Delta u = u^3 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

(c) Show that this solution is smooth.

Solution. Because $H_0^1(\Omega) \hookrightarrow L^4(\Omega)$, the infimum is positive. Let (u_m) be a minimizing sequence. It is bounded in $H_0^1(\Omega)$, so, after passing to a subsequence,

$$u_m \rightharpoonup u^* \quad \text{in } H_0^1(\Omega).$$

In dimension three the critical Sobolev exponent is 6. Rellich compactness therefore gives

$$H_0^1(\Omega) \Subset L^4(\Omega), \qquad u_m \longrightarrow u^* \quad \text{in } L^4(\Omega).$$

Consequently $\|u^*\|_{L^4(\Omega)} = 1$. By weak lower semicontinuity,

$$\int_\Omega |\nabla u^*|^2 \,\mathrm{d}x \le \liminf_{m\to\infty} \int_\Omega |\nabla u_m|^2 \,\mathrm{d}x = \lambda.$$

Since u^* is admissible, equality holds, and u^* attains the infimum.

For $v \in H_0^1(\Omega)$, consider

$$F(t) := \frac{1}{2} \frac{\displaystyle\int_\Omega |\nabla(u^* + tv)|^2 \,\mathrm{d}x}{\left(\displaystyle\int_\Omega |u^* + tv|^4 \,\mathrm{d}x\right)^{1/2}}.$$

The quotient is minimized at $t = 0$, and hence $F'(0) = 0$. Using $\|u^*\|_{L^4(\Omega)} = 1$ and $\int_\Omega |\nabla u^*|^2 \,\mathrm{d}x = \lambda$, we obtain

$$\int_\Omega \nabla u^* \cdot \nabla v \,\mathrm{d}x = \lambda \int_\Omega (u^*)^3 v \,\mathrm{d}x \qquad \text{for every } v \in H_0^1(\Omega).$$

Thus u^* is a weak solution of

$$-\Delta u^* = \lambda (u^*)^3, \qquad u^*\big|_{\partial\Omega} = 0.$$

Since $\lambda > 0$, the rescaled function

$$u := \sqrt{\lambda}\, u^*$$

satisfies

$$-\Delta u = u^3, \qquad u\big|_{\partial\Omega} = 0.$$

Finally, $u \in H_0^1(\Omega) \hookrightarrow L^6(\Omega)$, and so $u^3 \in L^2(\Omega)$. Dirichlet elliptic regularity yields $u \in H^2(\Omega)$. Since $H^2(\Omega)$ is an algebra in dimension three, $u^3 \in H^2(\Omega)$; applying elliptic regularity repeatedly gives $u \in H^k(\Omega)$ for every $k \ge 0$. Sobolev embedding then implies $u \in C^\infty(\overline{\Omega})$. □

Problem 10.2 (Spectral solutions of parabolic and dispersive equations)**.** Let $Q = (0,1)^2$, let $H = L^2(Q)$, and consider the Dirichlet operator

$$Av := \Delta v + a \cdot \nabla v + bv.$$

Assume that its Dirichlet realization has an orthonormal basis $(v_m)_{m\ge1}$ of H and eigenvalues $\lambda_m > 0$, with $\lambda_m \to \infty$, such that

$$Av_m = -\lambda_m v_m, \qquad v_m\big|_{\partial Q} = 0.$$

Given $f \in H$, do the following.

(a) Construct the solution of
$$\begin{cases} \partial_t u = Au & \text{in } Q \times (0, \infty), \\ u = 0 & \text{on } \partial Q \times (0, \infty), \\ u(\,\cdot\,, 0) = f & \text{in } Q, \end{cases}$$
and show that $u \in C([0, \infty); H)$ and is smooth in space for every $t > 0$.

(b) Construct the solution of
$$\begin{cases} \partial_t u = \mathrm{i}Au & \text{in } Q \times \mathbb{R}, \\ u = 0 & \text{on } \partial Q \times \mathbb{R}, \\ u(\,\cdot\,, 0) = f & \text{in } Q, \end{cases}$$
and show that $u \in C(\mathbb{R}; H)$.

Solution. Seeking a separated solution $u(x, y, t) = v(x, y)T(t)$ in the first equation gives
$$\frac{T'}{T} = \frac{\Delta v + a \cdot \nabla v + bv}{v} = -\lambda.$$
For the eigenpair (λ_m, v_m), therefore,
$$T_m'(t) + \lambda_m T_m(t) = 0, \qquad T_m(t) = \mathrm{e}^{-\lambda_m t}.$$
Expand the initial datum as
$$f = \sum_{m=1}^{\infty} b_m v_m, \qquad \sum_{m=1}^{\infty} |b_m|^2 < \infty.$$
The solution is
$$u(x, y, t) = \sum_{m=1}^{\infty} b_m v_m(x, y) \mathrm{e}^{-\lambda_m t}.$$
For every $t > 0$ and every $r \geq 0$,
$$\sum_{m=1}^{\infty} \lambda_m^{2r} |b_m|^2 \mathrm{e}^{-2\lambda_m t} \leq \left(\sup_{s>0} s^{2r} \mathrm{e}^{-2st} \right) \sum_{m=1}^{\infty} |b_m|^2 < \infty.$$
Elliptic regularity and Sobolev embedding therefore give spatial smoothness for $t > 0$.

Parseval's identity gives
$$\left\| u(\,\cdot\,, t) - f \right\|_H^2 = \sum_{m=1}^{\infty} |b_m|^2 \left| \mathrm{e}^{-\lambda_m t} - 1 \right|^2 \longrightarrow 0 \qquad (t \downarrow 0).$$
More generally, for $t \longrightarrow t_0 \geq 0$,
$$\| u(\,\cdot\,, t) - u(\,\cdot\,, t_0) \|_H^2 = \sum_{m=1}^{\infty} |b_m|^2 \left| \mathrm{e}^{-\lambda_m t} - \mathrm{e}^{-\lambda_m t_0} \right|^2 \longrightarrow 0.$$
Indeed, each summand converges to zero and is bounded by $4 |b_m|^2$, so dominated convergence applies. Hence $u \in C([0, \infty); H)$.

For the dispersive equation, the time factors are

$$T_m(t) = \mathrm{e}^{-\mathrm{i}\lambda_m t},$$

and therefore

$$u(x, y, t) = \sum_{m=1}^{\infty} b_m \mathrm{e}^{-\mathrm{i}\lambda_m t} v_m(x, y).$$

The series converges in H for every $t \in \mathbb{R}$, since Parseval gives

$$\|u(\,\cdot\,, t)\|_H^2 = \sum_{m=1}^{\infty} |b_m|^2 .$$

For $t \to t_0$, the same dominated-convergence argument applied to

$$\|u(\,\cdot\,, t) - u(\,\cdot\,, t_0)\|_H^2 = \sum_{m=1}^{\infty} |b_m|^2 \left|\mathrm{e}^{-\mathrm{i}\lambda_m t} - \mathrm{e}^{-\mathrm{i}\lambda_m t_0}\right|^2$$

shows that $u \in C(\mathbb{R}; H)$. □

Chapter 11

Local and Global Theory for Semilinear Heat Equations

11.1 Two questions for an evolution equation

For an evolution equation, there are two basic existence questions. First, can one construct a solution on a short time interval? This is the problem of *local existence*, for which Picard iteration and the contraction mapping principle are natural tools. Second, can the solution be continued for all positive times? This is the problem of *global existence*, and it generally requires energy estimates or other a priori bounds.

Consider, as a first example,

$$\partial_t u = \Delta u + |\nabla u|^2, \qquad x \in \mathbb{R}^3, \qquad u(0) = u_0. \tag{11.1}$$

To run a fixed-point argument, one must choose a Banach space in which the nonlinearity is controlled. In particular, if X is a Banach algebra and $\nabla u \in X$, then

$$\big\||\nabla u|^2\big\|_X \lesssim \|\nabla u\|_X^2.$$

The corresponding bound for higher integral powers follows by induction. The choice of space is therefore part of the argument, not merely a matter of notation.

11.1.1 The wrong and right iterations

A direct iteration would be

$$\partial_t u^{(k+1)} = \Delta u^{(k)} + |\nabla u^{(k)}|^2. \tag{11.2}$$

This scheme is unsuitable in a fixed Sobolev space. If $u^{(k)} \in H^s$, then $\Delta u^{(k)} \in H^{s-2}$, so two spatial derivatives are lost at each step.

Instead, keep the highest-order linear term at the new iterate:

$$\partial_t u^{(k+1)} = \Delta u^{(k+1)} + |\nabla u^{(k)}|^2. \tag{11.3}$$

The heat flow is then treated exactly. Since the nonlinearity loses only one derivative, the smoothing of the heat semigroup can recover that derivative.

11.2 Duhamel's formula and heat-flow smoothing

Consider the forced linear problem

$$\partial_t v = \Delta v + F, \qquad v(0) = v_0. \tag{11.4}$$

Duhamel's formula gives

$$v(t) = e^{t\Delta} v_0 + \int_0^t e^{(t-s)\Delta} F(s)\, \mathrm{d}s. \tag{11.5}$$

For $0 < t \leq 1$ and $\ell \geq 0$, the heat semigroup gains one derivative:

$$\|e^{t\Delta} f\|_{H^{\ell+1}} \lesssim t^{-1/2} \|f\|_{H^\ell}. \tag{11.6}$$

Consequently,

$$\|v(t)\|_{H^{\ell+1}} \lesssim \|v_0\|_{H^{\ell+1}} + \int_0^t (t-s)^{-1/2} \|F(s)\|_{H^\ell}\, \mathrm{d}s. \tag{11.7}$$

For equation (11.1), take $F = |\nabla v|^2$. When $\ell > 3/2$, the space $H^\ell(\mathbb{R}^3)$ is a Banach algebra, and hence

$$\big\||\nabla v|^2\big\|_{H^\ell} \lesssim \|\nabla v\|_{H^\ell}^2 \lesssim \|v\|_{H^{\ell+1}}^2. \tag{11.8}$$

The lecture chooses $\ell = 2$, so the iteration is carried out in $H^3(\mathbb{R}^3)$. Combining (11.7) and (11.8) yields

$$\begin{aligned} \|v(t)\|_{H^3} &\lesssim \|v_0\|_{H^3} + \int_0^t (t-s)^{-1/2} \|v(s)\|_{H^3}^2\, \mathrm{d}s \\ &\lesssim \|v_0\|_{H^3} + \sqrt{t} \sup_{0\leq s\leq t} \|v(s)\|_{H^3}^2. \end{aligned} \tag{11.9}$$

The integrable time singularity is precisely what makes the local argument possible.

11.3 Local existence for the gradient equation

Set

$$X_T^3 := C\big([0,T]; H^3(\mathbb{R}^3)\big), \qquad \|v\|_{X_T^3} := \sup_{0\leq t\leq T} \|v(t)\|_{H^3}.$$

Define

$$\Phi(v)(t) := e^{t\Delta} v_0 + \int_0^t e^{(t-s)\Delta} |\nabla v(s)|^2\, \mathrm{d}s. \tag{11.10}$$

11.3.1 Continuity in time

The heat semigroup is strongly continuous on every H^s. With the Fourier convention fixed earlier in the book,

$$\|e^{t\Delta} v_0 - v_0\|_{H^s}^2 = \int_{\mathbb{R}^3} (1+|\xi|^2)^s \big|e^{-4\pi^2 t|\xi|^2} - 1\big|^2 |\widehat{v}_0(\xi)|^2\, \mathrm{d}\xi \longrightarrow 0 \tag{11.11}$$

as $t \downarrow 0$, by dominated convergence.

For the Duhamel term, write $G(s) = |\nabla v(s)|^2$. If $v \in X_T^3$, then $G \in C([0,T]; H^2)$. For $t \geq t_0$, the semigroup property gives

$$\begin{aligned}\int_0^t e^{(t-s)\Delta} G(s)\ \mathrm{d}s - \int_0^{t_0} e^{(t_0-s)\Delta} G(s)\ \mathrm{d}s \\ = \int_{t_0}^t e^{(t-s)\Delta} G(s)\ \mathrm{d}s + (e^{(t-t_0)\Delta} - I) \int_0^{t_0} e^{(t_0-s)\Delta} G(s)\ \mathrm{d}s.\end{aligned}$$

The first term tends to zero in H^3 by (11.6), since it is bounded by

$$C\sqrt{t - t_0} \sup_{0 \leq s \leq T} \|G(s)\|_{H^2}.$$

The second tends to zero by strong continuity because the fixed integral belongs to H^3. The case $t < t_0$ is analogous. Thus Φ maps X_T^3 to itself.

11.3.2 The contraction

On an H^3-ball of radius R, the quadratic nonlinearity is locally Lipschitz:

$$\big\| |\nabla v|^2 - |\nabla w|^2 \big\|_{H^2} \leq C_R \|v - w\|_{H^3}. \tag{11.12}$$

Indeed, factor the difference into a sum of products and use the H^2 algebra estimate. It follows that

$$\|\Phi(v) - \Phi(w)\|_{X_T^3} \leq C_R \sqrt{T}\, \|v - w\|_{X_T^3}. \tag{11.13}$$

Take, for example, the closed ball of radius 2 about the constant path $t \mapsto v_0$. Strong continuity and (11.9) show that, after T is reduced, Φ maps this ball into itself. Reduce T once more so that $C_R\sqrt{T} < 1/2$. The iterates

$$v^{(k+1)} = \Phi(v^{(k)})$$

then remain in the ball and satisfy

$$\|v^{(k+1)} - v^{(k)}\|_{X_T^3} \leq \frac{1}{2} \|v^{(k)} - v^{(k-1)}\|_{X_T^3}.$$

The Banach fixed-point theorem gives the following conclusion.

Theorem 11.1 (Local existence for the gradient equation)**.** *If $v_0 \in H^3(\mathbb{R}^3)$, then for some $T > 0$ equation* (11.1) *has a unique mild solution*

$$v \in C\big([0,T]; H^3(\mathbb{R}^3)\big).$$

11.4 Continuation and global existence

Local existence alone does not guarantee a solution for all time. For an ordinary differential equation, global existence follows if the vector field is globally Lipschitz or if one has an a priori bound that prevents the solution from escaping every bounded set.

For example, suppose $\lambda \geq 0$ and

$$\ddot{x} + x + \lambda x^3 = 0. \tag{11.14}$$

Multiplication by $\dot{x}$ gives the conserved energy

$$\frac{1}{2}|\dot{x}(t)|^2 + \frac{1}{2}|x(t)|^2 + \frac{\lambda}{4}|x(t)|^4 = \frac{1}{2}|\dot{x}(0)|^2 + \frac{1}{2}|x(0)|^2 + \frac{\lambda}{4}|x(0)|^4. \tag{11.15}$$

This yields a uniform bound for both x and $\dot{x}$, and therefore global existence.

More generally, consider

$$\dot{x} = Ax + f(x), \qquad x(t_0) = x_{t_0}. \tag{11.16}$$

with f locally Lipschitz. The local fixed-point map based at t_0 is

$$\Phi_{t_0}(y)(t) = e^{A(t-t_0)}x_{t_0} + \int_{t_0}^{t} e^{A(t-s)} f(y(s))\, \mathrm{d}s. \tag{11.17}$$

If the solution remains uniformly bounded, the Lipschitz constant of f on the relevant ball is uniform. Hence the local lifespan can be chosen independently of the restart time t_0. Repeating the local construction excludes a finite maximal time.

The same continuation principle applies to a partial differential equation: a local solution can be continued as long as the norm used in its local theory remains finite.

11.5 The cubic semilinear heat equation

Consider

$$\partial_t u = \Delta u + \lambda u^3, \qquad x \in \mathbb{R}^3, \qquad u(0) = u_0. \tag{11.18}$$

Because $H^2(\mathbb{R}^3)$ is a Banach algebra, the map

$$\Psi(v)(t) := e^{t\Delta}u_0 + \lambda \int_0^t e^{(t-s)\Delta} v(s)^3\, \mathrm{d}s \tag{11.19}$$

is a contraction on a suitable ball in $C([0,T]; H^2(\mathbb{R}^3))$ when T is small. Thus every $u_0 \in H^2$ gives a unique local solution. To make this solution global by the preceding continuation principle, one would initially seek a time-independent H^2 bound.

11.6 The multiplier method and a priori estimates

The lecture organizes energy estimates under the mnemonic of testing the terms

$$A\partial_t u + B \cdot \nabla u + Cu$$

against appropriate multipliers. For (11.18), the first two useful multipliers are u and $\partial_t u$. We first make the calculation for smooth solutions that decay at spatial infinity; equivalently, one may integrate over balls and let their radii tend to infinity. The identities extend to the mild solutions by approximation.

11.6.1 Multiplication by the solution

Multiplying (11.18) by u gives

$$\partial_t\left(\frac{u^2}{2}\right) = u\,\Delta u + \lambda u^4$$
$$= \partial_i\big(u\,\partial_i u\big) - |\nabla u|^2 + \lambda u^4. \tag{11.20}$$

After integration in space and time,

$$\boxed{\frac{1}{2}\|u(t)\|_{L^2}^2 + \int_0^t \left(\|\nabla u(s)\|_{L^2}^2 - \lambda\|u(s)\|_{L^4}^4\right) \mathrm{d}s = \frac{1}{2}\|u_0\|_{L^2}^2.} \tag{11.21}$$

For the defocusing sign $\lambda = -1$, all terms on the left are nonnegative. Therefore

$$u \in L_t^\infty L_x^2, \qquad \nabla u \in L_t^2 L_x^2, \qquad u \in L_t^4 L_x^4, \tag{11.22}$$

with norms controlled by the initial L^2 norm.

11.6.2 Multiplication by the time derivative

Multiplying (11.18) by $\partial_t u$ gives the pointwise identity

$$|\partial_t u|^2 = \partial_i\big((\partial_t u)\partial_i u\big) - \frac{1}{2}\partial_t|\nabla u|^2 + \frac{\lambda}{4}\partial_t|u|^4. \tag{11.23}$$

Integration yields

$$\boxed{\begin{aligned}&\int_0^t \|\partial_s u(s)\|_{L^2}^2\,\mathrm{d}s + \frac{1}{2}\|\nabla u(t)\|_{L^2}^2 - \frac{\lambda}{4}\|u(t)\|_{L^4}^4\\ &\quad = \frac{1}{2}\|\nabla u_0\|_{L^2}^2 - \frac{\lambda}{4}\|u_0\|_{L^4}^4.\end{aligned}} \tag{11.24}$$

Here $\partial_s u(s)$ denotes the time derivative evaluated at time s.

In three dimensions, the Gagliardo–Nirenberg inequality gives

$$\|u\|_{L^4} \lesssim \|\nabla u\|_{L^2}^{3/4}\|u\|_{L^2}^{1/4}. \tag{11.25}$$

Thus initial data in H^1, and in particular the H^2 data used in the first local construction, have finite energy.

For $\lambda = -1$, combining (11.21) and (11.24) gives the a priori bounds

$$\begin{aligned}\sup_{0\le s\le t}\|u(s)\|_{H^1}^2 &+ \int_0^t \|\nabla u(s)\|_{L^2}^2\,\mathrm{d}s + \int_0^t \|\partial_s u(s)\|_{L^2}^2\,\mathrm{d}s\\ &+ \int_0^t \|u(s)\|_{L^4}^4\,\mathrm{d}s \le C\Big(\|u_0\|_{H^1}^2 + \|u_0\|_{H^1}^4\Big).\end{aligned} \tag{11.26}$$

These estimates are uniform in time, but they control only H^1 rather than the H^2 norm used in the first local theory.

11.7 Lowering the local theory to H^1

For the defocusing equation

$$\partial_t u = \Delta u - u^3, \qquad x \in \mathbb{R}^3, \tag{11.27}$$

Duhamel's formula is

$$u(t) = e^{t\Delta} u_0 - \int_0^t e^{(t-s)\Delta} u(s)^3 \, \mathrm{d}s. \tag{11.28}$$

A direct H^1 estimate on the differentiated nonlinearity does not close: $H^1(\mathbb{R}^3)$ is not a Banach algebra, and $\|\nabla(u^3)\|_{L^2}$ cannot be controlled solely by $\|u\|_{H^1}^3$. The remedy is to sacrifice a factor $(t-s)^{-1/2}$ in time and let the heat flow supply the spatial derivative:

$$\begin{aligned} \|e^{(t-s)\Delta} u(s)^3\|_{H^1} &\lesssim (t-s)^{-1/2} \|u(s)^3\|_{L^2} \\ &= (t-s)^{-1/2} \|u(s)\|_{L^6}^3 \\ &\lesssim (t-s)^{-1/2} \|u(s)\|_{H^1}^3. \end{aligned} \tag{11.29}$$

Thus, with $X_T^1 = C([0,T]; H^1)$,

$$\|u\|_{X_T^1} \lesssim \|u_0\|_{H^1} + \sqrt{T}\, \|u\|_{X_T^1}^3. \tag{11.30}$$

Moreover, on an H^1-ball of radius R,

$$\|u^3 - v^3\|_{L^2} \le C_R \|u - v\|_{H^1}, \tag{11.31}$$

because $H^1(\mathbb{R}^3) \hookrightarrow L^6(\mathbb{R}^3)$. The same heat-flow estimate then makes the Duhamel map a contraction when T is small. We obtain local existence in $C([0,T]; H^1)$.

The a priori estimate (11.26) now controls exactly the norm used by this local theory. The continuation argument therefore gives a global solution.

Theorem 11.2 (Global defocusing cubic heat flow)**.** *For every $u_0 \in H^1(\mathbb{R}^3)$, equation (11.27) has a unique global mild solution*

$$u \in C\big([0,\infty); H^1(\mathbb{R}^3)\big).$$

It satisfies the a priori estimates (11.26).

11.8 Gronwall's inequality and propagation of regularity

We recall the elementary form of Gronwall's inequality used to propagate higher regularity.

Theorem 11.3 (Gronwall)**.** *Let $x\colon [0,T] \to [0,\infty)$ be absolutely continuous. Suppose $b \ge 0$ is integrable and*

$$\dot{x}(t) \le b(t) x(t)$$

for almost every t. Then

$$x(t) \le x(0) \exp\left(\int_0^t b(s) \, \mathrm{d}s \right), \qquad 0 \le t \le T. \tag{11.32}$$

Proof. The integrating factor gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\int_0^t b(s)\,\mathrm{d}s}x(t)\right) = e^{-\int_0^t b(s)\,\mathrm{d}s}\big(\dot{x}(t) - b(t)x(t)\big) \le 0.$$

Evaluating at 0 proves (11.32). □

11.8.1 The H^2 estimate

Suppose first that $u_0 \in H^2(\mathbb{R}^3)$. Using one derivative of heat smoothing in (11.28), and working on a fixed interval $[0, T]$, we have

$$\|u(t)\|_{H^2} \lesssim \|u_0\|_{H^2} + C_T \int_0^t (t-s)^{-1/2}\|u(s)^3\|_{H^1}\,\mathrm{d}s. \tag{11.33}$$

Here the global semigroup bound contains $1 + (t-s)^{-1/2}$; on $[0, T]$ this is bounded by a constant depending on T times $(t-s)^{-1/2}$. The product is estimated by

$$\begin{aligned}\|u^3\|_{L^2} &= \|u\|_{L^6}^3 \lesssim \|u\|_{H^1}^3,\\ \|\nabla(u^3)\|_{L^2} &\lesssim \|u\|_{L^6}^2\|\nabla u\|_{L^6} \lesssim \|u\|_{H^1}^2\|u\|_{H^2}.\end{aligned}$$

Let

$$M := 1 + \sup_{s\ge 0}\|u(s)\|_{H^1}, \qquad y(t) := \|u(t)\|_{H^2}.$$

The global H^1 estimate makes M finite, and

$$y(t) \le Cy(0) + C_T M^2 \int_0^t (t-s)^{-1/2} y(s)\,\mathrm{d}s. \tag{11.34}$$

To reduce this to ordinary Gronwall, substitute the inequality into its own convolution. Since

$$\int_0^t (t-s)^{-1/2}s^{-1/2}\,\mathrm{d}s = \pi,$$

one obtains, on every fixed interval $[0, T]$,

$$y(t) \le C_T y(0) + CM^4 \int_0^t y(s)\,\mathrm{d}s.$$

Gronwall's inequality therefore bounds y on every finite time interval. Hence

$$u_0 \in H^2(\mathbb{R}^3) \quad\Longrightarrow\quad u \in C\big([0,\infty); H^2(\mathbb{R}^3)\big). \tag{11.35}$$

11.8.2 Higher Sobolev regularity

For every integer $k \ge 2$, $H^k(\mathbb{R}^3)$ is a Banach algebra. More precisely, the polynomial nonlinearity satisfies the tame estimate

$$\|u^3\|_{H^k} \lesssim \|u\|_{H^2}^2\|u\|_{H^k}. \tag{11.36}$$

The heat semigroup is contractive on H^k, so Duhamel's formula gives

$$\|u(t)\|_{H^k} \lesssim \|u_0\|_{H^k} + \int_0^t \|u(s)\|_{H^2}^2 \|u(s)\|_{H^k} \, \mathrm{d}s. \tag{11.37}$$

The H^2 norm is bounded on every finite interval, and Gronwall's inequality implies

$$u_0 \in H^k(\mathbb{R}^3), \quad k \geq 2 \quad \Longrightarrow \quad u \in C\big([0,\infty); H^k(\mathbb{R}^3)\big). \tag{11.38}$$

Summary

The defocusing cubic heat equation is globally well posed in $H^1(\mathbb{R}^3)$ because the energy estimate controls the same norm used in the local fixed-point theory. Once global existence is known at that level, heat smoothing and Gronwall estimates propagate every higher integer Sobolev regularity present in the initial data.

Chapter 12

Nonlinear Evolution Equations: Local Theory, Energy Estimates, and Finite Propagation

12.1 Local theory and global continuation

For nonlinear evolution equations, the basic plan has two parts.

1. Prove local existence by Picard iteration, using the contraction mapping principle in a suitable function space.
2. Obtain global existence from an a priori estimate that prevents the solution norm from becoming infinite in finite time.

The function space used for Picard iteration must interact well with the nonlinearity. For powers of a function, this often means choosing a Sobolev space that is an algebra. Derivative nonlinearities require an additional ingredient because the nonlinearity may apparently lose derivatives.

12.2 A derivative nonlinearity

Consider

$$\partial_t u = \Delta u + |\nabla u|^2, \qquad x \in \mathbb{R}^3, \qquad u(0) = u_0. \tag{12.1}$$

The Picard iterates solve

$$\partial_t u^{(n+1)} = \Delta u^{(n+1)} + |\nabla u^{(n)}|^2, \qquad u^{(n+1)}(0) = u_0. \tag{12.2}$$

It is important that the Laplacian acts on the new iterate $u^{(n+1)}$. Replacing it by $\Delta u^{(n)}$ would lose two derivatives at every step.

To estimate an iterate, first consider the linear problem

$$\partial_t v = \Delta v + F, \qquad v(0) = v_0. \tag{12.3}$$

Duhamel's principle gives

$$v(t) = \mathrm{e}^{t\Delta} v_0 + \int_0^t \mathrm{e}^{(t-s)\Delta} F(s)\, \mathrm{d}s. \tag{12.4}$$

For $0 < t - s \le 1$, the heat semigroup gains one derivative with the estimate

$$\left\| \mathrm{e}^{(t-s)\Delta} F(s) \right\|_{H^{\ell+1}} \le \frac{C}{(t-s)^{1/2}} \|F(s)\|_{H^\ell}. \tag{12.5}$$

If $\ell > 3/2$, then $H^\ell(\mathbb{R}^3)$ is an algebra. Hence

$$\left\| |\nabla u|^2 \right\|_{H^\ell} \le C \|u\|_{H^{\ell+1}}^2. \tag{12.6}$$

Combining (12.4)–(12.6) gives, for $T \le 1$,

$$\|u^{(n+1)}\|_{C([0,T];H^{\ell+1})} \le \|u_0\|_{H^{\ell+1}} + CT^{1/2} \|u^{(n)}\|_{C([0,T];H^{\ell+1})}^2. \tag{12.7}$$

The corresponding difference estimate makes the Picard map a contraction on a sufficiently small ball. Taking $\ell = 2$ yields local existence in $H^3(\mathbb{R}^3)$. This argument gives a local solution; a separate a priori estimate is needed for global existence.

12.3 Continuation and a finite-dimensional model

For comparison, consider the ordinary differential equation

$$\dot{x} = Ax + f(x), \qquad x(t_0) = x_0, \tag{12.8}$$

where f is locally Lipschitz. The variation-of-constants formula is

$$x(t) = \mathrm{e}^{A(t-t_0)} x(t_0) + \int_{t_0}^t \mathrm{e}^{A(t-s)} f(x(s))\, \mathrm{d}s. \tag{12.9}$$

If $|x(t)|$ remains bounded, then the local Lipschitz and size bounds for f remain uniform on the ball containing the trajectory. Picard's argument can therefore be restarted with a time interval whose length depends only on that ball. Repeating this step continues the solution.

An energy or Lyapunov function can provide the required bound. For example, consider

$$\ddot{x} + x + \lambda x^3 = 0, \qquad \lambda \ge 0. \tag{12.10}$$

Multiplication by $\dot{x}$ gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{2} \dot{x}^2 + \frac{1}{2} x^2 + \frac{\lambda}{4} x^4 \right) = 0. \tag{12.11}$$

Thus

$$\frac{1}{2}\dot{x}(t)^2 + \frac{1}{2}x(t)^2 + \frac{\lambda}{4}x(t)^4 = \frac{1}{2}\dot{x}(0)^2 + \frac{1}{2}x(0)^2 + \frac{\lambda}{4}x(0)^4. \tag{12.12}$$

The right-hand side bounds both x and $\dot{x}$, so the continuation argument gives a global solution.

The same principle applies to a PDE: if its local theory is constructed in H^3, then a time-independent bound for the H^3 norm permits the local solution to be restarted and continued globally.

12.4 The cubic heat equation

Consider the semilinear equation

$$\partial_t u = \Delta u + \lambda u^3, \qquad x \in \mathbb{R}^3, \qquad u(0) = u_0. \tag{12.13}$$

Since $2 > 3/2$, the space $H^2(\mathbb{R}^3)$ is an algebra. The iterates

$$\partial_t u^{(n+1)} = \Delta u^{(n+1)} + \lambda \left(u^{(n)}\right)^3, \qquad u^{(n+1)}(0) = u_0, \tag{12.14}$$

therefore converge on a sufficiently short interval when $u_0 \in H^2$. This proves local existence. To use continuation at this regularity, one would need an a priori H^2 bound independent of the length of the time interval.

12.4.1 Two energy identities

Assume first that the solution is smooth and decays at spatial infinity. A useful general multiplier ansatz for an evolution equation is

$$A\,\partial_t u + B \cdot \nabla u + Cu.$$

The coefficients are chosen to produce a favorable energy identity. For (12.13), the first useful choice is simply the multiplier u.

Multiplying (12.13) by u gives the pointwise identity

$$\partial_t\left(\frac{1}{2}u^2\right) = \partial_i(u\,\partial_i u) - |\nabla u|^2 + \lambda|u|^4.$$

After integration over $\mathbb{R}^3$, the divergence term vanishes and

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)\|_{L^2}^2 + \|\nabla u(t)\|_{L^2}^2 - \lambda\|u(t)\|_{L^4}^4 = 0. \tag{12.15}$$

In the dissipative case $\lambda = -1$, integration in time yields

$$\frac{1}{2}\|u(t)\|_{L^2}^2 + \int_0^t \|\nabla u(s)\|_{L^2}^2\,\mathrm{d}s + \int_0^t \|u(s)\|_{L^4}^4\,\mathrm{d}s = \frac{1}{2}\|u_0\|_{L^2}^2. \tag{12.16}$$

Multiplying the equation instead by $\partial_t u$ gives

$$|\partial_t u|^2 = \partial_i(\partial_i u\,\partial_t u) - \partial_t\left(\frac{1}{2}|\nabla u|^2\right) + \frac{\lambda}{4}\partial_t(u^4).$$

Consequently, when $\lambda = -1$,

$$\begin{aligned}\frac{1}{2}\|\nabla u(t)\|_{L^2}^2 + \frac{1}{4}\|u(t)\|_{L^4}^4 + \int_0^t \|\partial_s u(s)\|_{L^2}^2\,\mathrm{d}s \\ = \frac{1}{2}\|\nabla u_0\|_{L^2}^2 + \frac{1}{4}\|u_0\|_{L^4}^4.\end{aligned} \tag{12.17}$$

Together, (12.16) and (12.17) give, on every finite time interval,

$$\|u\|_{L_t^\infty H_x^1} \leq C(\|u_0\|_{H^1}),$$
$$\|\nabla u\|_{L_t^2 L_x^2} \leq C(\|u_0\|_{H^1}),$$
$$\|\partial_t u\|_{L_t^2 L_x^2} \leq C(\|u_0\|_{H^1}).$$

Match the local theory to the energy

These estimates do not directly control the H^2 norm used in the first Picard construction. They do, however, control H^1. This suggests rebuilding the local theory at the energy regularity and then using the same estimates for global continuation.

12.4.2 Local and global theory in H^1

For $\lambda = -1$, the mild form of (12.13) is

$$u(t) = \mathrm{e}^{t\Delta} u_0 - \int_0^t \mathrm{e}^{(t-s)\Delta} u(s)^3 \, \mathrm{d}s. \tag{12.18}$$

The one-derivative heat estimate gives

$$\left\|\mathrm{e}^{(t-s)\Delta} u(s)^3\right\|_{H^1} \leq \frac{C}{(t-s)^{1/2}} \|u(s)^3\|_{L^2}. \tag{12.19}$$

Sobolev embedding in three dimensions implies

$$\|u^3\|_{L^2} = \|u\|_{L^6}^3 \leq C\|\nabla u\|_{L^2}^3 \leq C\|u\|_{H^1}^3. \tag{12.20}$$

Therefore

$$\|u\|_{C([0,T];H^1)} \leq C\|u_0\|_{H^1} + CT^{1/2}\|u\|_{C([0,T];H^1)}^3. \tag{12.21}$$

The analogous difference estimate makes the mild map a contraction for small T. Thus one has local existence and uniqueness in H^1. The uniform H^1 estimate above prevents the local norm from blowing up, and the continuation argument gives a global H^1 solution.

The energy identities also suggest a weaker compactness construction at energy regularity. By themselves, however, such estimates do not provide a uniqueness theory; uniqueness here comes from the mild contraction argument.

12.4.3 Persistence of higher regularity

Suppose now that $u_0 \in H^k(\mathbb{R}^3)$ for an integer $k \geq 1$. The case $k = 2$ shows the main point. From (12.18) and the one-derivative gain,

$$\|u(t)\|_{H^2} \leq C\|u_0\|_{H^2} + C\int_0^t \frac{\|u(s)^3\|_{H^1}}{(t-s)^{1/2}} \, \mathrm{d}s. \tag{12.22}$$

Using $H^1(\mathbb{R}^3) \hookrightarrow L^6(\mathbb{R}^3)$,

$$\begin{aligned} \|u^3\|_{H^1} &\leq C\left(\|u\|_{L^6}^3 + \|u\|_{L^6}^2\|\nabla u\|_{L^6}\right) \\ &\leq C\|u\|_{H^1}^2\|u\|_{H^2}. \end{aligned} \tag{12.23}$$

The global H^1 bound thus reduces (12.22) on $[0, T]$ to

$$X(t) \leq a + b\int_0^t \frac{X(s)}{(t-s)^{1/2}}\,\mathrm{d}s, \qquad X(t) = \|u(t)\|_{H^2}, \tag{12.24}$$

with finite constants a and b depending on the data and T.

For comparison, the ordinary Gronwall inequality begins with

$$x(t) \leq a + \lambda\int_0^t x(s)\,\mathrm{d}s.$$

If $z(t) = a + \lambda\int_0^t x(s)\,\mathrm{d}s$, then $\dot{z} = \lambda x \leq \lambda z$, and hence

$$x(t) \leq z(t) \leq a\mathrm{e}^{\lambda t}. \tag{12.25}$$

The same proof works with a nonnegative integrable coefficient in place of λ: if

$$x(t) \leq a + \int_0^t b(s)x(s)\,\mathrm{d}s,$$

then

$$x(t) \leq a\exp\left(\int_0^t b(s)\,\mathrm{d}s\right).$$

To handle (12.24), apply the inequality to itself once. Since

$$\int_s^t \frac{\mathrm{d}r}{(t-r)^{1/2}(r-s)^{1/2}} = \pi, \tag{12.26}$$

one obtains an ordinary Gronwall inequality for X. Therefore X stays finite on every bounded time interval. Repeating the derivative estimate gives persistence of every higher Sobolev regularity H^k present in the initial data.

12.5 The cubic wave equation and finite propagation

Consider the defocusing cubic wave equation on $\mathbb{R}^3$,

$$\partial_t^2 u - \Delta u + u^3 = 0. \tag{12.27}$$

Its energy space is $H^1 \times L^2$; by comparison, the linear wave equation propagates data in $H^{k+1} \times H^k$.

Multiplication of (12.27) by $\partial_t u$ gives the local conservation law

$$\partial_t\left(\frac{1}{2}|\partial_t u|^2 + \frac{1}{2}|\nabla u|^2 + \frac{1}{4}u^4\right) - \partial_i(\partial_i u\,\partial_t u) = 0. \tag{12.28}$$

After integration over all of space,

$$E(t) := \int_{\mathbb{R}^3} \left(\frac{1}{2}|\partial_t u|^2 + \frac{1}{2}|\nabla u|^2 + \frac{1}{4}u^4 \right) \mathrm{d}x = E(0). \tag{12.29}$$

The energy does not contain the L^2 norm of u. On a finite interval it can be recovered from

$$u(t) = u(0) + \int_0^t \partial_s u(s)\ \mathrm{d}s. \tag{12.30}$$

More generally, let $m = (m^0, m^1, m^2, m^3)$ be the outward spacetime unit normal to a hypersurface. The flux associated with (12.28) is

$$m^0 e(u) - m^i \partial_i u\, \partial_t u.$$

When $m^0 \geq |(m^1, m^2, m^3)|$ and $m^0 > 0$, this expression is nonnegative after completing squares. The limiting case of equality occurs on a light cone.

Fix x_0, a time $t > 0$, and a radius $R > 0$, and consider the truncated backward cone

$$K = \{(s, x) : 0 \leq s \leq t,\ |x - x_0| \leq R + t - s\}. \tag{12.31}$$

Its top is $T = \{t\} \times B_R(x_0)$, its base is $B = \{0\} \times B_{R+t}(x_0)$, and its lateral boundary is denoted by L.

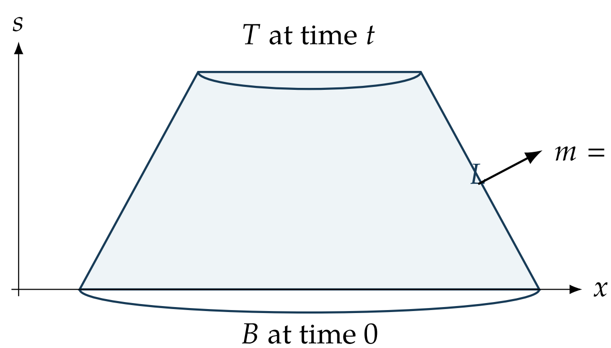


Figure 12.1: A truncated backward light cone. Its lateral normal has equal time and radial components.

Write

$$e(u) = \frac{1}{2}|\partial_t u|^2 + \frac{1}{2}|\nabla u|^2 + \frac{1}{4}u^4.$$

The spacetime divergence theorem applied to (12.28) gives

$$\int_T e(u)\ \mathrm{d}x + \int_L \frac{1}{\sqrt{2}} (e(u) - \partial_r u\, \partial_t u)\ \mathrm{d}S = \int_B e(u)\ \mathrm{d}x. \tag{12.32}$$

On the lateral boundary,

$$e(u) - \partial_r u\, \partial_t u = \frac{1}{2}|\partial_t u - \partial_r u|^2 + \frac{1}{2}|\nabla_T u|^2 + \frac{1}{4}u^4 \geq 0, \tag{12.33}$$

where $\nabla_T u$ is the angular part of the spatial gradient. It follows that

$$\int_{B_R(x_0)} e(u)(t,x)\,\mathrm{d}x \le \int_{B_{R+t}(x_0)} e(u)(0,x)\,\mathrm{d}x. \tag{12.34}$$

Thus the energy in the top of the cone is controlled entirely by the initial energy in its base. This is the energy form of finite propagation speed.

12.6 A scalar convection–diffusion model

As a model for the cancellation in the Navier–Stokes nonlinearity, consider the scalar equation on $\mathbb{R}^2$

$$\partial_t u - \Delta u = \partial_x(u^2). \tag{12.35}$$

Multiplication by u gives

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)\|_{L^2}^2 + \|\nabla u(t)\|_{L^2}^2 = \int_{\mathbb{R}^2} u\,\partial_x(u^2)\,\mathrm{d}x. \tag{12.36}$$

For decaying functions, the nonlinear term drops out because

$$\int_{\mathbb{R}^2} u\,\partial_x(u^2)\,\mathrm{d}x = \frac{2}{3}\int_{\mathbb{R}^2} \partial_x(u^3)\,\mathrm{d}x = 0. \tag{12.37}$$

Hence

$$\frac{1}{2}\|u(t)\|_{L^2}^2 + \int_0^t \|\nabla u(s)\|_{L^2}^2\,\mathrm{d}s = \frac{1}{2}\|u_0\|_{L^2}^2. \tag{12.38}$$

This a priori estimate is the starting point for constructing a global solution from L^2 data by approximation and compactness.

12.7 Sobolev interpolation and nonlinear estimates

The nonlinear estimates used above fit into a general interpolation principle. Let $u : \mathbb{R}^d \to \mathbb{R}$, let $0 \le j < k$, and let $1 \le p, q, r \le \infty$. Suppose α satisfies

$$\frac{1}{r} = \frac{j}{d} + \alpha\left(\frac{1}{p} - \frac{k}{d}\right) + \frac{1-\alpha}{q}, \qquad \frac{j}{k} \le \alpha \le 1. \tag{12.39}$$

Then the Gagliardo–Nirenberg interpolation inequality has the form

$$\|D^j u\|_{L^r} \le C\|D^k u\|_{L^p}^{\alpha}\|u\|_{L^q}^{1-\alpha}. \tag{12.40}$$

As an important example, suppose $k > d/p$ and $u \in W^{k,p}(\mathbb{R}^d)$. Set $j = 0$ and $q = p$. The right-hand side of (12.39) is positive at $\alpha = 0$ and negative at $\alpha = 1$, so one can choose

$$\alpha = \frac{d}{kp} \in (0,1) \qquad \text{and} \qquad r = \infty.$$

Therefore

$$\|u\|_{L^\infty} \le C\|D^k u\|_{L^p}^{d/(kp)}\|u\|_{L^p}^{1-d/(kp)} \le C\|u\|_{W^{k,p}}. \tag{12.41}$$

This embedding yields the product and composition estimates needed for nonlinear PDE. For an integer $k > d/p$,

$$\|uv\|_{W^{k,p}} \le C\|u\|_{W^{k,p}}\|v\|_{W^{k,p}}, \tag{12.42}$$

so $W^{k,p}(\mathbb{R}^d)$ is an algebra. In particular, for every positive integer m,

$$\|D^k(u^m)\|_{L^p} \le C_m\|u\|_{W^{k,p}}^m, \qquad \|u^m\|_{W^{k,p}} \le C_m\|u\|_{W^{k,p}}^m. \tag{12.43}$$

Finally, let $F \in C^N(\mathbb{R})$ with $N \ge k$ and $F(0) = 0$. The chain rule, Hölder's inequality, and Sobolev embedding give a composition estimate of the form

$$\|F(u)\|_{W^{k,p}} \le C_{F,\|u\|_{L^\infty}}\left(\|u\|_{W^{k,p}} + \|u\|_{W^{k,p}}^k\right). \tag{12.44}$$

The constant depends only on the derivatives of F over the bounded range of u. The condition $F(0) = 0$ ensures that $F(u)$ is integrable on the whole space; it is unnecessary on a bounded spatial domain.

12.8 Problems

Problem 12.1. Let $k \in \mathbb{N}$, $1 \le p \le \infty$, and $k > d/p$. For an integer $j \ge 0$, write $\partial^j u$ for the family $\{\partial^\alpha u : |\alpha| = j\}$, with the corresponding summed L^p-norm.

Suppose that $F \in C^N(\mathbb{R}, \mathbb{R})$, $N > k$, and $F(0) = 0$. Prove the Sobolev algebra and composition estimates below. In the last estimate, $m := \|u\|_{L^\infty}$, and the constant may depend on F and m.

(a) For $u, v \in \mathcal{S}(\mathbb{R}^d)$ and exponents satisfying

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2},$$

show that

$$\left\|\partial_j(uv)\right\|_{L^p} \lesssim \left\|\partial_j u\right\|_{L^{p_1}}\|v\|_{L^{p_2}} + \|u\|_{L^{p_1}}\left\|\partial_j v\right\|_{L^{p_2}}.$$

(b) Using the interpolation inequality

$$\left\|\partial^i f\right\|_{L^r} \lesssim \left\|f\right\|_{L^r}^{1-\theta}\left\|\partial^k f\right\|_{L^r}^{\theta}, \qquad \theta = \frac{i}{k} \le 1,$$

show that

$$\left\|\partial^k(uv)\right\|_{L^p} \lesssim \left\|\partial^k u\right\|_{L^{p_1}}\|v\|_{L^{p_2}} + \|u\|_{L^{p_1}}\left\|\partial^k v\right\|_{L^{p_2}}.$$

(c) Show that, if $u \in W^{k,p}(\mathbb{R}^d)$, then

$$\left\|\partial^k(u^2)\right\|_{L^p} \lesssim \|u\|_{W^{k,p}}^2.$$

(d) Show by induction that, for every positive integer M,

$$\left\|u^M\right\|_{W^{k,p}} \lesssim \|u\|_{W^{k,p}}^M .$$

(e) Conclude that

$$\|F(u)\|_{W^{k,p}} \le C_{F,m}\big(\|u\|_{W^{k,p}} + \|u\|_{W^{k,p}}^k\big).$$

Solution. For part (a), the product rule followed by Hölder's inequality gives

$$\left\|\partial_j(uv)\right\|_{L^p} \le \left\|(\partial_j u)v\right\|_{L^p} + \left\|u\,\partial_j v\right\|_{L^p} \le \left\|\partial_j u\right\|_{L^{p_1}} \|v\|_{L^{p_2}} + \|u\|_{L^{p_1}} \left\|\partial_j v\right\|_{L^{p_2}} .$$

For part (b), the Leibniz rule yields

$$\left\|\partial^k(uv)\right\|_{L^p} \lesssim \left\|\partial^k u\right\|_{L^{p_1}} \|v\|_{L^{p_2}} + \|u\|_{L^{p_1}} \left\|\partial^k v\right\|_{L^{p_2}} + \sum_{j=1}^{k-1} \left\|\partial^j u\right\|_{L^{p_1}} \left\|\partial^{k-j} v\right\|_{L^{p_2}} .$$

For $1 \le j \le k-1$, interpolation bounds the corresponding mixed term by

$$\begin{aligned}
&\left\|\partial^j u\right\|_{L^{p_1}} \left\|\partial^{k-j} v\right\|_{L^{p_2}} \\
&\quad\lesssim \left\|\partial^k u\right\|_{L^{p_1}}^{j/k} \|u\|_{L^{p_1}}^{1-j/k} \left\|\partial^k v\right\|_{L^{p_2}}^{1-j/k} \|v\|_{L^{p_2}}^{j/k} \\
&\quad= \big(\left\|\partial^k u\right\|_{L^{p_1}} \|v\|_{L^{p_2}}\big)^{j/k} \big(\|u\|_{L^{p_1}} \left\|\partial^k v\right\|_{L^{p_2}}\big)^{1-j/k} \\
&\quad\lesssim \left\|\partial^k u\right\|_{L^{p_1}} \|v\|_{L^{p_2}} + \|u\|_{L^{p_1}} \left\|\partial^k v\right\|_{L^{p_2}} ,
\end{aligned}$$

where the final step is Young's product inequality. Summing over j proves the claim.

For part (c), the Gagliardo–Nirenberg–Sobolev interpolation inequality, with the L^∞-norm as the lower-order endpoint, gives

$$\left\|\partial^j u\right\|_{L^{kp/j}} \lesssim \left\|\partial^k u\right\|_{L^p}^{j/k} \|u\|_{L^\infty}^{1-j/k} , \qquad 1 \le j \le k.$$

Consequently, Hölder's inequality with exponents kp/j and $kp/(k-j)$, followed by Young's product inequality, gives

$$\left\|\partial^k(uv)\right\|_{L^p} \lesssim \left\|\partial^k u\right\|_{L^p} \|v\|_{L^\infty} + \|u\|_{L^\infty} \left\|\partial^k v\right\|_{L^p} .$$

Taking $v = u$ and using $W^{k,p}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ therefore gives

$$\left\|\partial^k(u^2)\right\|_{L^p} \lesssim 2\,\|u\|_{L^\infty} \left\|\partial^k u\right\|_{L^p} \lesssim \|u\|_{W^{k,p}}^2 .$$

For part (d), the assertion is immediate when $M = 1$. Assume it holds with $M-1$ in place of M. Applying the preceding product estimate to $u\,u^{M-1}$ gives

$$\left\|\partial^k(u^M)\right\|_{L^p} \lesssim \left\|\partial^k u\right\|_{L^p} \left\|u^{M-1}\right\|_{L^\infty} + \|u\|_{L^\infty} \left\|\partial^k(u^{M-1})\right\|_{L^p}$$
$$\lesssim \|u\|_{W^{k,p}}^M .$$

The same calculation for each order $0 \le j \le k$, followed by summation, yields

$$\left\|u^M\right\|_{W^{k,p}} \lesssim \|u\|_{W^{k,p}}^M .$$

For part (e), the scanned solution applies Taylor's theorem in the form

$$F(x) = F'(0)x + \cdots + \frac{F^{(k-1)}(0)}{(k-1)!}x^{k-1} + \frac{F^{(k)}(\xi_x)}{k!}x^k,$$

where ξ_x lies between 0 and x, and then invokes part (d) term by term to obtain

$$\|F(u)\|_{W^{k,p}} \le C_{F,m}\big(\|u\|_{W^{k,p}} + \|u\|_{W^{k,p}}^k\big).$$

The displayed Taylor calculation records the argument in the source. Its remainder coefficient depends on the spatial point after $x = u(x)$ is substituted, so a complete justification of the final step uses the standard Sobolev composition estimate rather than treating that coefficient as a constant. □

Problem 12.2. Let $f \in C^1(\mathbb{R}, \mathbb{R})$ satisfy $f(0) = 0$, and let $u_0 \in H^1(\mathbb{R})$. Consider

$$\partial_t u = \partial_x^2 u + f(u), \qquad u(0) = u_0.$$

(a) Show that $\mathrm{e}^{t\partial_x^2} : H^1(\mathbb{R}) \to H^1(\mathbb{R})$ and

$$\left\|\mathrm{e}^{t\partial_x^2}\right\|_{H^1 \to H^1} \le 1.$$

(b) Show that, for $u \in C([0,T], H^1(\mathbb{R}))$,

$$\int_0^t \mathrm{e}^{(t-s)\partial_x^2} f(u(s))\,\mathrm{d}s \in C([0,T], H^1(\mathbb{R})).$$

(c) For sufficiently small $T > 0$, study the iteration

$$u^{n+1}(t) = \mathrm{e}^{t\partial_x^2} u_0 + \int_0^t \mathrm{e}^{(t-s)\partial_x^2} f(u^n(s))\,\mathrm{d}s, \qquad u^0(t) \equiv 0,$$

Assume the local Lipschitz estimate stated in the solution. Show that the iteration converges in $C([0,T], H^1(\mathbb{R}))$.

(d) Conclude that the fixed point is the unique mild solution in the class on which that Lipschitz estimate holds.

Solution. For part (a), the Fourier transform gives

$$\begin{aligned}\left\| e^{t\partial_x^2} u_0 \right\|_{H^1}^2 &= \int_{\mathbb{R}} (1+|\xi|^2) e^{-8\pi^2 t|\xi|^2} |\widehat{u_0}(\xi)|^2 \, d\xi \\ &\le \|u_0\|_{H^1}^2 ,\end{aligned}$$

because $0 < e^{-8\pi^2 t|\xi|^2} \le 1$. Here the factor $4\pi^2$ in the heat multiplier follows from the Fourier convention fixed earlier in the book.

For part (b), let $S(t) = e^{t\partial_x^2}$ and $g(s) = f(u(s))$. The chain rule, $f(0) = 0$, and the embedding $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$ show that $g(s) \in H^1(\mathbb{R})$. They also show that $g \in C([0,T], H^1(\mathbb{R}))$, and hence that g is bounded there. Define

$$\Phi(u)(t) = \int_0^t S(t-s) g(s) \, ds.$$

For $h > 0$, the semigroup property gives

$$\begin{aligned}\|\Phi(u)(t+h) - \Phi(u)(t)\|_{H^1} \le &\int_t^{t+h} \left\| S(t+h-s) g(s) \right\|_{H^1} \, ds \\ &+ \int_0^t \left\| (S(h) - I) S(t-s) g(s) \right\|_{H^1} \, ds.\end{aligned}$$

The first term tends to zero because g is bounded. The second tends to zero by strong continuity of the heat semigroup and the dominated convergence theorem. Hence $\Phi(u) \in C([0,T], H^1)$.

For part (c), the handwritten argument uses a constant K such that

$$\left\| f(v) - f(w) \right\|_{H^1} \le K \left\| v - w \right\|_{H^1}$$

on a bounded set containing all iterates. With

$$d_n(T) := \sup_{0 \le t \le T} \left\| u^{n+1}(t) - u^n(t) \right\|_{H^1} ,$$

the semigroup bound from part (a) gives

$$\begin{aligned}\left\| u^{n+1}(t) - u^n(t) \right\|_{H^1} &\le \int_0^t \left\| f(u^n(s)) - f(u^{n-1}(s)) \right\|_{H^1} \, ds \\ &\le KT \sup_{0 \le s \le T} \left\| u^n(s) - u^{n-1}(s) \right\|_{H^1} .\end{aligned}$$

Thus $d_n(T) \le KT d_{n-1}(T)$. Choose T so that $KT \le 1/4$. Since

$$d_0(T) = \sup_{0 \le t \le T} \left\| e^{t\partial_x^2} u_0 \right\|_{H^1} \le \|u_0\|_{H^1},$$

one obtains

$$d_n(T) \le 4^{-n} \|u_0\|_{H^1}.$$

Therefore (u^n) is Cauchy in the Banach space $C([0,T], H^1)$, and its limit u satisfies

$$u(t) = e^{t\partial_x^2} u_0 + \int_0^t e^{(t-s)\partial_x^2} f(u(s))\, ds.$$

For part (d), the same contraction estimate shows that two fixed points in the chosen set agree. Thus the iteration and its geometric difference bound are precisely the contraction-mapping argument recorded in the source.

The local H^1-Lipschitz estimate used in part (c) is not automatic for an arbitrary $f \in C^1(\mathbb{R})$. For this particular contraction proof, one needs an additional condition such as local Lipschitz continuity of f' (for example, $f \in C^2$), or a different existence argument. □

Problem 12.3. Let $S(t) = e^{t\Delta}$ be the heat semigroup on $\mathbb{R}^d$, and let $\|S(t)\|_{X \to Y}$ denote its operator norm from X to Y. Establish the following bounds.

(a) For $s \in \mathbb{R}$,

$$\|S(t)\|_{H^s \to H^s} \le 1.$$

(b) For $k \in \mathbb{N}$ and $1 \le p \le \infty$,

$$\|S(t)\|_{W^{k,p} \to W^{k,p}} \lesssim 1.$$

(c) If $1 \le p \le q \le \infty$, then

$$\|S(t)\|_{L^p \to L^q} \lesssim t^{-\frac{d}{2}(\frac{1}{p} - \frac{1}{q})}.$$

(d) For $1 \le p \le \infty$, show that

$$\|S(t)\|_{L^p \to W^{1,p}} \lesssim 1 + t^{-1/2}.$$

In particular, the right-hand side is $\lesssim t^{-1/2}$ when $0 < t \le 1$.

Solution. Write

$$K_t(x) = \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{4t}\right), \qquad S(t)u = K_t * u.$$

For part (a),

$$\|S(t)u\|_{H^s}^2 = \int_{\mathbb{R}^d} (1 + |\xi|^2)^s e^{-8\pi^2 t|\xi|^2} |\widehat{u}(\xi)|^2\, d\xi$$

$$\leq \|u\|_{H^s}^2 .$$

For part (b), derivatives commute with convolution, so Young's inequality and $\|K_t\|_{L^1} = 1$ give

$$\begin{aligned} \|S(t)u\|_{W^{k,p}} &\lesssim \sum_{|\alpha|\leq k} \|K_t * \partial^\alpha u\|_{L^p} \\ &\leq \sum_{|\alpha|\leq k} \|K_t\|_{L^1} \|\partial^\alpha u\|_{L^p} \lesssim \|u\|_{W^{k,p}} . \end{aligned}$$

For part (c), choose r so that

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

Young's inequality and a change of variables yield

$$\begin{aligned} \|S(t)u\|_{L^q} &\leq \|K_t\|_{L^r} \|u\|_{L^p} , \\ \|K_t\|_{L^r} &= C_{d,r}\, t^{-\frac{d}{2}(1-1/r)} = C_{d,p,q}\, t^{-\frac{d}{2}(1/p-1/q)}. \end{aligned}$$

This proves the claimed L^p-L^q estimate.

For part (d), Young's inequality gives

$$\begin{aligned} \|S(t)u\|_{W^{1,p}} &\lesssim \left(\|K_t\|_{L^1} + \|\nabla K_t\|_{L^1}\right) \|u\|_{L^p} \\ &\lesssim (1 + t^{-1/2}) \|u\|_{L^p} , \end{aligned}$$

because $\partial_j K_t(x) = -(x_j/(2t))K_t(x)$ and hence $\|\nabla K_t\|_{L^1} \lesssim t^{-1/2}$.

These estimates apply, for example, to the semilinear problem

$$\partial_t u = \Delta u + u^2 \quad \text{on } \mathbb{R}^2, \qquad u(0) = u_0 \in L^2(\mathbb{R}^2).$$

Indeed, the bound $S(t) : L^1 \to L^2$ has size $Ct^{-1/2}$, which is the heat estimate used in a Picard iteration for this equation. □

Problem 12.4. Let $u_0 \in H^3(\mathbb{R}^2)$, and consider the viscous Burgers equation

$$\partial_t u = \Delta u + u\, \partial_1 u = \Delta u + \frac{1}{2}\partial_1(u^2), \qquad u(0) = u_0.$$

First establish local well-posedness in $C([0,T], H^3(\mathbb{R}^2))$. Then:

(a) derive the L^2-energy identity;
(b) derive the first-derivative energy identity;

(c) estimate its cubic term by Sobolev interpolation and Young's inequality;
(d) deduce a uniform H^1-bound;
(e) obtain the corresponding second-derivative estimate and conclude global well-posedness;
(f) explain how the argument propagates higher spatial regularity and time differentiability.

Solution. For local existence, define

$$\Phi(v)(t) = \mathrm{e}^{t\Delta} u_0 + \frac{1}{2} \int_0^t \mathrm{e}^{(t-s)\Delta} \partial_1 \big(v(s)^2\big) \, \mathrm{d}s.$$

The heat-semigroup estimate, followed by the algebra estimate in $H^3(\mathbb{R}^2)$, gives

$$\begin{aligned} \|\Phi(v)(t)\|_{H^3} &\le \|u_0\|_{H^3} + C \int_0^t (t-s)^{-1/2} \left\|v(s)^2\right\|_{H^3} \, \mathrm{d}s \\ &\le \|u_0\|_{H^3} + C \int_0^t (t-s)^{-1/2} \|v(s)\|_{H^3}^2 \, \mathrm{d}s. \end{aligned}$$

With

$$\|v\|_{X_T} := \sup_{0 \le t \le T} \|v(t)\|_{H^3},$$

this becomes

$$\|\Phi(v)\|_{X_T} \le \|u_0\|_{H^3} + C\sqrt{T} \, \|v\|_{X_T}^2.$$

The analogous difference estimate is

$$\|\Phi(v) - \Phi(w)\|_{X_T} \le C\sqrt{T}\big(\|v\|_{X_T} + \|w\|_{X_T}\big) \|v - w\|_{X_T}.$$

Thus, for $T > 0$ sufficiently small, Φ is a contraction on a closed ball in $C([0,T], H^3(\mathbb{R}^2))$. Banach's fixed-point theorem gives a unique local solution, and the Duhamel formula shows that it satisfies both the equation and the initial condition.

For the energy estimates, all integrations below are over $\mathbb{R}^2$. Boundary terms vanish by approximation from rapidly decaying functions.

(a) Multiplication by u gives

$$\partial_t \frac{u^2}{2} = u\Delta u + u^2 \partial_1 u = \operatorname{div}(u\nabla u) - |\nabla u|^2 + \frac{1}{3}\partial_1(u^3).$$

After integration,

$$\frac{1}{2} \|u(t)\|_{L^2}^2 + \int_0^t \|\nabla u(s)\|_{L^2}^2 \, \mathrm{d}s = \frac{1}{2} \|u_0\|_{L^2}^2.$$

(b) Put $\partial_i u := \partial_{x_i} u$. Differentiating the equation gives

$$\partial_t \partial_i u = \Delta \partial_i u + (\partial_i u)\partial_1 u + u \, \partial_{1i} u.$$

Multiplying by $\partial_i u$, summing in i, and integrating by parts, we obtain

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\nabla u\|_{L^2}^2 + \left\|D^2 u\right\|_{L^2}^2 = \int |\nabla u|^2 \, \partial_1 u \, \mathrm{d}x + \int u \, \nabla u \cdot \partial_1 \nabla u \, \mathrm{d}x$$

$$= \frac{1}{2}\int |\nabla u|^2\, \partial_1 u \ \mathrm{d}x.$$

Consequently,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\, \|\nabla u\|_{L^2}^2 + \left\|D^2 u\right\|_{L^2}^2 \le C \int |\nabla u|^3 \ \mathrm{d}x.$$

(c) Hölder's inequality and the two-dimensional Sobolev interpolation inequality give

$$\begin{aligned}\int |\nabla u|^3 \ \mathrm{d}x &\le \|\nabla u\|_{L^2}\, \|\nabla u\|_{L^4}^2 \\ &\le C\, \|\nabla u\|_{L^2}^2 \left\|D^2 u\right\|_{L^2} \\ &\le \frac{1}{2}\left\|D^2 u\right\|_{L^2}^2 + C\, \|\nabla u\|_{L^2}^4 .\end{aligned}$$

Hence

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\, \|\nabla u\|_{L^2}^2 + \frac{1}{2}\left\|D^2 u\right\|_{L^2}^2 \le C\, \|\nabla u\|_{L^2}^4 .$$

(d) If $Y(t) = \|\nabla u(t)\|_{L^2}^2$, then

$$Y'(t) \le C Y(t)^2.$$

Part (a) supplies the additional integral bound

$$\int_0^t Y(s)\ \mathrm{d}s \le \frac{1}{2}\, \|u_0\|_{L^2}^2 .$$

Writing the differential inequality as $Y' \le (CY)Y$ and applying Gronwall's inequality yields

$$Y(t) \le Y(0) \exp\!\left(C \int_0^t Y(s)\ \mathrm{d}s\right) \le Y(0) \exp\!\left(C\, \|u_0\|_{L^2}^2\right).$$

Thus $\|u(t)\|_{H^1}$ is bounded uniformly in time. Integrating the estimate from part (c) and using this uniform bound together with part (a) also gives

$$\int_0^t \left\|D^2 u(s)\right\|_{L^2}^2 \ \mathrm{d}s \le C(u_0)$$

for every $t > 0$.

(e) Differentiate the equation by ∂_{ij}. Then

$$\begin{aligned}\partial_t \partial_{ij} u - \Delta \partial_{ij} u &= \partial_{ij}(u \partial_1 u) \\ &= (\partial_{ij} u)\partial_1 u + (\partial_i u)\partial_{1j} u + (\partial_j u)\partial_{1i} u + u\, \partial_{1ij} u.\end{aligned}$$

Multiplication by $\partial_{ij} u$, summation in i, j, and integration by parts leave terms bounded by

$$C \int \left|D^2 u\right|^2 |\nabla u| \ \mathrm{d}x.$$

Therefore

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|D^2 u\right\|_{L^2}^2 + \left\|D^3 u\right\|_{L^2}^2 \le C \int \left|D^2 u\right|^2 |\nabla u| \ \mathrm{d}x.$$

Using Hölder's inequality and, again, the two-dimensional Sobolev interpolation inequality,

$$\begin{aligned}\int \left|D^2u\right|^2 |\nabla u| \, \mathrm{d}x &\leq \left\|D^2u\right\|_{L^4}^2 \|\nabla u\|_{L^2} \\ &\leq C \left\|D^3u\right\|_{L^2} \left\|D^2u\right\|_{L^2} \|\nabla u\|_{L^2} \\ &\leq \frac{1}{2} \left\|D^3u\right\|_{L^2}^2 + C \left\|D^2u\right\|_{L^2}^2 \|\nabla u\|_{L^2}^2 .\end{aligned}$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\|D^2u\right\|_{L^2}^2 + \left\|D^3u\right\|_{L^2}^2 \leq C \|\nabla u\|_{L^2}^2 \left\|D^2u\right\|_{L^2}^2 .$$

Since $\|\nabla u\|_{L^2}^2$ is integrable in time by part (a), Gronwall's inequality yields

$$\left\|D^2u(t)\right\|_{L^2}^2 \leq \left\|D^2u_0\right\|_{L^2}^2 \exp\left(C \int_0^t \|\nabla u(s)\|_{L^2}^2 \, \mathrm{d}s\right) \leq C(u_0).$$

Thus the H^2-norm cannot blow up in finite time. Repeating the same energy argument for higher spatial derivatives gives global continuation in $H^\ell(\mathbb{R}^2)$ whenever the corresponding initial norm is finite.

(f) Time differentiability follows directly from the equation. For $j = 0$ there is nothing to prove. If

$$\partial_t^j u \in C\big([0,T], H^{\ell-2j}(\mathbb{R}^2)\big), \qquad \ell \geq 2j,$$

then

$$\partial_t^{j+1} u = \Delta \partial_t^j u + \frac{1}{2} \partial_1 \partial_t^j (u^2),$$

and the product estimates used above give the next level of regularity. This closes the induction. □

Problem 12.5. Let $2 \leq p < 5$, and let $u \in H^1(\mathbb{R}^3)$ solve

$$-\Delta u + u = u^p \qquad \text{in } \mathbb{R}^3.$$

Show that $u \in H^2(\mathbb{R}^3)$. If $p \in \{2,3,4\}$, bootstrap the equation to prove that $u \in C^\infty(\mathbb{R}^3)$.

Solution. When $2 \leq p \leq 3$, Sobolev interpolation gives $H^1(\mathbb{R}^3) \hookrightarrow L^{2p}(\mathbb{R}^3)$. Hence

$$\|u^p\|_{L^2} = \|u\|_{L^{2p}}^p < \infty,$$

and elliptic regularity for $-\Delta + I$ yields $u \in H^2(\mathbb{R}^3)$.

Now suppose $3 < p < 5$. Since $u \in L^6$, initially $u^p \in L^{6/p}$. Starting with $r_0 = 6$, suppose inductively that $u \in L^{r_j}$ and set $q_j = r_j/p$. As long as $1 < q_j < 3/2$, elliptic L^{q_j}-regularity and Sobolev embedding give

$$u \in W^{2,q_j}(\mathbb{R}^3) \hookrightarrow L^{r_{j+1}}(\mathbb{R}^3), \qquad \frac{1}{r_{j+1}} = \frac{p}{r_j} - \frac{2}{3}.$$

If $q_j \geq 3/2$, the critical or supercritical Sobolev embedding instead gives $u \in L^r$ for every finite r, and in particular $u \in L^{2p}$.

To see that the iteration reaches this range, write $a_j = 1/r_j$. While the displayed recurrence applies,

$$a_{j+1} = pa_j - \frac{2}{3}, \qquad a_j - \frac{2}{3(p-1)} = p^j\left(\frac{1}{6} - \frac{2}{3(p-1)}\right).$$

Because $p < 5$, the quantity in parentheses is negative. Thus after finitely many steps either $q_j \geq 3/2$ or $r_j \geq 2p$. In both cases $u \in L^{2p}$, so $u^p \in L^2$, and elliptic regularity yields $u \in H^2(\mathbb{R}^3)$.

The interpolation calculation recorded in the scan now gives the useful a priori estimate. The relation

$$\frac{1}{2p} = \left(\frac{1}{6} - \frac{1}{3}\right)\alpha + \frac{1-\alpha}{6}$$

gives $\alpha = (p-3)/(2p)$. Therefore

$$\begin{aligned} \|u\|_{H^2} &\leq C\,\|u^p\|_{L^2} = C\,\|u\|_{L^{2p}}^p \\ &\leq C\,\left\|D^2 u\right\|_{L^2}^{(p-3)/2}\,\|u\|_{L^6}^{(p+3)/2} \\ &\leq C\big(\|u\|_{H^1}\big)\,\|u\|_{H^2}^{(p-3)/2}. \end{aligned}$$

Because $0 < (p-3)/2 < 1$, this inequality gives an H^2-bound. Thus $u \in H^2(\mathbb{R}^3)$ throughout the range $2 \leq p < 5$.

Since $2 > 3/2$, Sobolev embedding now gives $u \in L^\infty(\mathbb{R}^3)$. Differentiating the equation,

$$-\Delta\partial_i u + \partial_i u = pu^{p-1}\partial_i u.$$

Consequently,

$$\|\partial_i u\|_{H^2} \leq C\left\|pu^{p-1}\partial_i u\right\|_{L^2} \leq C\left\|u^{p-1}\right\|_{L^\infty}\|\partial_i u\|_{L^2},$$

so $u \in H^3(\mathbb{R}^3)$. For $p \in \{2,3,4\}$, $H^k(\mathbb{R}^3)$ is an algebra once $k \geq 2$. If $u \in H^k$, then $u^p \in H^k$, and elliptic regularity gives

$$\|u\|_{H^{k+2}} \leq C\,\|u^p\|_{H^k} \leq C\big(\|u\|_{H^k}\big).$$

Induction yields $u \in H^k(\mathbb{R}^3)$ for every k, and hence $u \in C^\infty(\mathbb{R}^3)$. □

Problem 12.6. Consider the wave equation on $\mathbb{R}^2$,

$$\partial_t^2 u - \Delta u = 0, \qquad u(x,0) = f(x) \in H^1_{\text{loc}}(\mathbb{R}^2), \qquad \partial_t u(x,0) = g(x) \in L^2_{\text{loc}}(\mathbb{R}^2).$$

Establish existence in the local energy class. Then prove the energy inequality on a truncated spacetime cone of slope one and use it to show finite propagation and uniqueness.

Solution. Let R denote the forward fundamental solution of the wave operator in $\mathbb{R}^{1+2}$. The representation

$$u = \partial_t R * f + R * g$$

gives a solution satisfying

$$u \in C\big([0,\infty), H^1_{\text{loc}}(\mathbb{R}^2)\big) \cap C^1\big([0,\infty), L^2_{\text{loc}}(\mathbb{R}^2)\big).$$

For the energy estimate, multiply the equation by $\partial_t u$. Since

$$0 = \partial_t u\big(\partial_t^2 u - \Delta u\big)$$

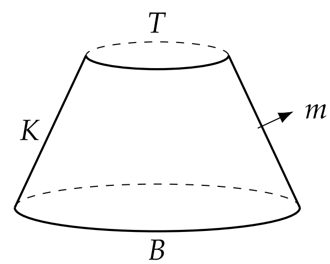
$$= \partial_t \left(\frac{|\partial_t u|^2}{2} + \frac{|\nabla_x u|^2}{2} \right) - \sum_{j=1}^{2} \partial_j \big(\partial_j u \, \partial_t u \big),$$

the spacetime vector field

$$P = \left(\frac{|\partial_t u|^2 + |\nabla_x u|^2}{2}, -\partial_t u \, \nabla_x u \right)$$

is divergence free.

Let Ω be a truncated cone whose lower disk is $B \subset \{t = 0\}$, whose upper disk is T, and whose lateral null boundary is K:



Here $m = (m^t, m_x)$ is the outward unit normal. The divergence theorem gives

$$0 = \int_{\partial\Omega} m \cdot P \; \mathrm{d}S.$$

On B, $m^t = -1$ and $m_x = 0$; on T, $m^t = 1$ and $m_x = 0$. Along the lateral boundary K, the cone has slope one, so

$$m \cdot P = m^t \frac{|\partial_t u|^2 + |\nabla_x u|^2}{2} - m_x \cdot \nabla_x u \, \partial_t u \geq 0,$$

by Cauchy's inequality. It follows that

$$\int_T \left(\frac{|\partial_t u|^2}{2} + \frac{|\nabla_x u|^2}{2} \right) \mathrm{d}S_T \leq \int_B \left(\frac{|\partial_t u|^2}{2} + \frac{|\nabla_x u|^2}{2} \right) \mathrm{d}S_B.$$

Thus vanishing Cauchy data on B force the solution to vanish on every later section T of the cone. This is finite propagation with speed one.

Finally, if u_1 and u_2 are two local-energy solutions with the same initial data, their difference $w = u_1 - u_2$ has

$$w|_B = 0, \qquad \partial_t w|_B = 0.$$

The cone inequality gives zero energy on every later section. Hence $\partial_t w = 0$ and $\nabla_x w = 0$ there; together with the zero trace on B, this implies $w = 0$. Therefore the solution is unique in

$$C\big([0,\infty), H^1_{\mathrm{loc}}(\mathbb{R}^2)\big) \cap C^1\big([0,\infty), L^2_{\mathrm{loc}}(\mathbb{R}^2)\big).$$

□

Chapter 13

Energy Estimates, Nonlinear Iteration, and Hyperbolic Uniqueness

The opening roadmap for this lecture lists energy estimates, hyperbolic equations, uniqueness, and potential theory. The surviving pages develop the first three themes: how an a priori estimate controls an iteration, how to repair derivative loss, and how uniqueness follows either from energy or from analytic duality. No potential-theory calculation follows in this scan.

13.1 What a nonlinear energy inequality gives

A useful estimate is not merely an inequality: its shape must force the unknown quantity to remain in a bounded region. Let $X \geq 0$ denote an energy or a space–time norm.

13.1.1 A superlinear error term

Suppose that

$$X \leq CX^{1+\delta} + A, \qquad C > 0, \quad \delta > 0. \tag{13.1}$$

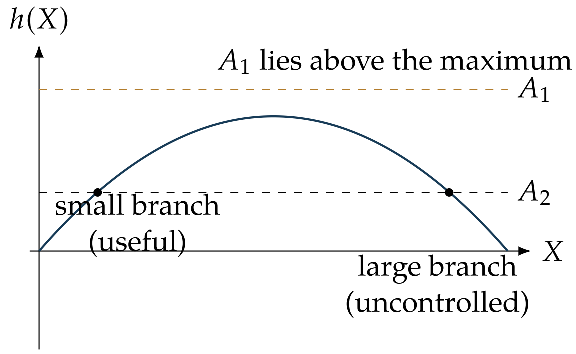
Equivalently,

$$h(X) := X - CX^{1+\delta} \leq A.$$

The graph of h first rises and then tends to $-\infty$. Its unique positive maximum occurs at

$$X_* = \big(C(1+\delta)\big)^{-1/\delta}.$$

If A lies below $h(X_*)$, then the equation $h(X) = A$ has a small root and a large root. Inequality (13.1) alone allows both branches, so the large branch is useless. If, however, $X = X(T)$ depends continuously on time, begins on the small branch, and the estimate holds for every intermediate time, then $X(T)$ cannot cross the forbidden interval between the two roots. This is the mechanism behind a local estimate.



For example, an energy may have the form

$$E(T) = \int_0^T \int_{\mathbb{R}} |\partial_x u(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t, \qquad E(0) = 0.$$

An estimate schematically resembling

$$\|u(T)\| \le A + T\|u\|_{L^2}^2$$

keeps the solution on the good branch when T is sufficiently small. Small A, small C, or a small time factor in front of the nonlinear term can all produce this situation. Large A or large C need not give any bound.

The model inequality

$$X \le A + TX^2 \tag{13.2}$$

makes the bootstrap explicit. Assume temporarily that $X \le 2A$. If $T < 1/(4A)$, then

$$A + TX^2 \le A + 4TA^2 < 2A.$$

The strict improvement closes the assumption by continuity and gives $X \le 2A$ on the chosen time interval.

13.1.2 A sublinear error term

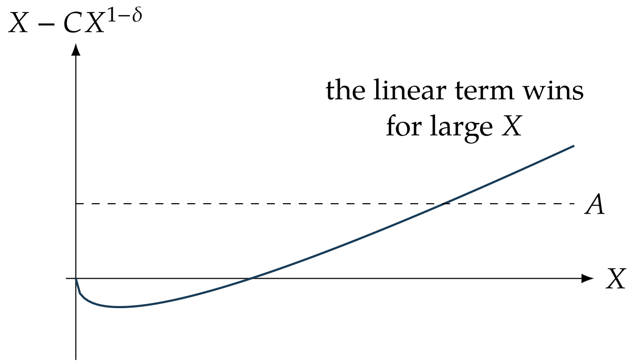
Now suppose

$$X \le CX^{1-\delta} + A. \tag{13.3}$$

For every $\delta > 0$, the nonlinear term grows more slowly than X for large X. Indeed, if $X \ge (2C)^{1/\delta}$, then $CX^{1-\delta} \le X/2$, and hence

$$X \le \max\left\{(2C)^{1/\delta}, 2A\right\}. \tag{13.4}$$

Unlike the superlinear estimate, this gives a global bound without choosing a small time. At the endpoint $\delta = 0$, one instead needs $C < 1$, in which case $X \le A/(1-C)$.



13.2 A priori bounds for a semilinear heat equation

Consider the Cauchy problem on the line

$$\partial_t u = \partial_x^2 u + f(u), \qquad u(0) = u_0 \in H^1(\mathbb{R}). \tag{13.5}$$

Assume $f \in C^1(\mathbb{R})$ and $f(0) = 0$. The latter condition ensures that $f(u)$ can belong to $L^2(\mathbb{R})$ when $u \in H^1(\mathbb{R})$. The Duhamel iteration is

$$u^{n+1}(t) = \mathrm{e}^{t\Delta} u_0 + \int_0^t \mathrm{e}^{(t-s)\Delta} f\big(u^n(s)\big)\, \mathrm{d}s. \tag{13.6}$$

An a priori estimate first gives a uniform bound for the sequence; a separate estimate must then give contraction.

For $T > 0$, set

$$\|u\|_T := \sup_{0 \le t \le T} \|u(t)\|_{H^1(\mathbb{R})}.$$

The heat semigroup is bounded on H^1, so

$$\|u^{n+1}(t)\|_{H^1} \le \|u_0\|_{H^1} + \int_0^t \big\|f(u^n(s))\big\|_{H^1}\, \mathrm{d}s. \tag{13.7}$$

In one dimension, $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$. If $\|u\|_{L^\infty} \le M$, then the mean-value theorem and the chain rule give

$$\|f(u)\|_{L^2} \le C_M \|u\|_{L^2}, \tag{13.8}$$
$$\|\partial_x(f(u))\|_{L^2} = \|f'(u)\partial_x u\|_{L^2} \le C_M \|\partial_x u\|_{L^2}, \tag{13.9}$$

where $C_M = \sup_{|z| \le M} |f'(z)|$. The model example in the notes is $f(u) = u^3$, for which

$$|u(x)^3| \le \|u\|_{L^\infty}^2 |u(x)|.$$

Thus (13.7) yields the schematic estimate

$$\|u^{n+1}\|_T \le A + T C_{\|u^n\|_T} \|u^n\|_T, \qquad A := \|u_0\|_{H^1}, \tag{13.10}$$

where the Sobolev-embedding constant is absorbed into C_R.

The dependence of C_R on the unknown norm makes this estimate circular unless it is used as a bootstrap. Assume $\|u^n\|_T \le 10A$ and choose T so that

$$TC_{10A} \le \delta_0 < \frac{9}{10}.$$

Then

$$\|u^{n+1}\|_T \le A + 10\delta_0 A < 10A.$$

Starting with an iterate in this ball, induction gives the uniform a priori bound

$$\|u^n\|_T < 10A \qquad (n \ge 1). \tag{13.11}$$

This is the iteration version of the small-branch argument.

13.3 Contraction and the loss of a derivative

Subtracting two consecutive Duhamel iterates gives

$$\|u^{n+1} - u^n\|_T \le T\big\|f(u^n) - f(u^{n-1})\big\|_{C([0,T];H^1)}. \tag{13.12}$$

The L^2 part on the right is controlled by the mean-value theorem. For the derivative part, however,

$$\begin{aligned}&\partial_x\big(f(u^n) - f(u^{n-1})\big)\\ &\quad = f'(u^n)\big(\partial_x u^n - \partial_x u^{n-1}\big) + \big(f'(u^n) - f'(u^{n-1})\big)\partial_x u^{n-1}.\end{aligned}$$

The first term is bounded on the uniform ball (13.11). A linear bound for the second term would require f' to be locally Lipschitz; the hypothesis $f \in C^1$ alone does not provide it. A direct H^1 contraction therefore asks for more regularity than the a priori estimate.

The heat flow repairs the derivative loss by gaining one derivative at the cost of a time singularity:

$$\|\mathrm{e}^{(t-s)\Delta} g\|_{H^1} \le \frac{C}{\sqrt{t-s}}\|g\|_{L^2}, \qquad 0 < t - s \le 1. \tag{13.13}$$

On the bounded iteration ball, f is Lipschitz, so

$$\|f(u^n) - f(u^{n-1})\|_{L^2} \le C_{10A}\|u^n - u^{n-1}\|_{L^2}.$$

Using (13.13) directly in the Duhamel formula gives

$$\begin{aligned}\|u^{n+1} - u^n\|_T &\le CC_{10A} \sup_{0\le t\le T} \int_0^t \frac{\mathrm{d}s}{\sqrt{t-s}} \|u^n - u^{n-1}\|_T\\ &\le 2CC_{10A}\sqrt{T}\,\|u^n - u^{n-1}\|_T. \end{aligned}\tag{13.14}$$

Taking T smaller once more makes the last factor less than one. This proves contraction without assuming that f' is Lipschitz.

There is a second useful strategy: keep the sequence uniformly bounded in the strong norm H^1, but prove contraction only in the weaker norm L^2. Since the heat semigroup is an L^2 contraction,

$$\|u^{n+1} - u^n\|_{C([0,T];L^2)} \le TC_{10A}\|u^n - u^{n-1}\|_{C([0,T];L^2)}. \tag{13.15}$$

Thus one obtains boundedness in the strongest available norm and convergence in a weaker norm.

13.4 Quasilinear iteration and an elliptic change of variable

The same issue appears when the nonlinearity contains derivatives, for example

$$\partial_t u = \partial_x^2 u + f(\partial_x u). \tag{13.16}$$

For a quasilinear wave equation, energy estimates replace heat smoothing. A model from the notes is

$$\partial_t^2 u - \partial_x\left[\left(1 + \frac{1}{2}\cos u\right)\partial_x u\right] = 0. \tag{13.17}$$

Because $1 + \frac{1}{2}\cos u \geq \frac{1}{2}$, the frozen-coefficient problem remains strictly hyperbolic. One may iterate by solving

$$\partial_t^2 u^{n+1} - \partial_x\left[\left(1 + \frac{1}{2}\cos u^n\right)\partial_x u^{n+1}\right] = 0, \tag{13.18}$$

and obtain an a priori bound by the same integration-by-parts calculation used for the linear wave equation.

The derivative difficulty is more severe for the fully nonlinear example

$$\partial_t^2 u - \partial_x^2 u + \varepsilon\left(\partial_x^2 u\right)^2 = 0. \tag{13.19}$$

Differentiation in x gives

$$\partial_x\partial_t^2 u - \partial_x^3 u + 2\varepsilon(\partial_x^2 u)(\partial_x^3 u) = 0. \tag{13.20}$$

Freezing one factor suggests the formal Newton–Picard equation

$$\partial_x\partial_t^2 u^{n+1} - \partial_x^3 u^{n+1} + 2\varepsilon(\partial_x^2 u^n)\partial_x^3 u^{n+1} = 0. \tag{13.21}$$

Its leading coefficient contains $\partial_x^2 u^n$. But controlling (13.20) first controls $\partial_x u$, and differentiation is not invertible: it loses the constant mode.

The remedy in the notes is to use an invertible elliptic operator. Set

$$v = (I - \partial_x^2)u, \qquad u = (I - \partial_x^2)^{-1}v. \tag{13.22}$$

With the Fourier normalization used in this book, the multiplier of $I - \partial_x^2$ on $\mathbb{R}$ is $1 + 4\pi^2\xi^2$, so its inverse is well defined on Sobolev spaces and gains two derivatives. Applying $I - \partial_x^2$ to (13.19), and using

$$\partial_x^4 u = \partial_x^2 u - \partial_x^2 v,$$

gives the exact transformed equation

$$\partial_t^2 v - \partial_x^2 v + 2\varepsilon(\partial_x^2 u)\partial_x^2 v = \varepsilon(\partial_x^2 u)^2 + 2\varepsilon(\partial_x^3 u)^2. \tag{13.23}$$

Consequently, an iteration may freeze $u = (I - \partial_x^2)^{-1}v$ at the previous step:

$$\begin{aligned}\partial_t^2 v^{n+1} - \partial_x^2 v^{n+1} &+ 2\varepsilon(\partial_x^2 u^n)\partial_x^2 v^{n+1} \\ &= \varepsilon(\partial_x^2 u^n)^2 + 2\varepsilon(\partial_x^3 u^n)^2.\end{aligned} \tag{13.24}$$

Provided $1 - 2\varepsilon\partial_x^2 u^n$ stays uniformly positive, this is a variable-coefficient hyperbolic equation for v^{n+1}. Energy estimates for v therefore produce a priori control in a high Sobolev norm, while the elliptic inverse recovers u without losing information.

13.5 Uniqueness through the adjoint equation

Energy is not the only route to uniqueness. Let L be a linear differential operator of order m with real-analytic coefficients, and let the analytic hypersurface B be non-characteristic for L. The Cauchy–Kowalevski theorem gives a unique analytic solution for analytic Cauchy data. Holmgren's theorem strengthens this conclusion to weak solutions. In its distributional form, a solution that vanishes on one side of B must vanish near B; in a class where the Cauchy traces are defined, this gives local uniqueness for zero Cauchy data.

For an order-m equation, the Cauchy problem prescribes the normal jet

$$\partial_\nu^j u\big|_B = f_j, \qquad j = 0, \dots, m-1.$$

The model in the notes is

$$\partial_t^2 u - a(x,t)\partial_x^2 u = 0, \qquad a(x,t) = 1 + \frac{1}{2}\cos(x^2 + t^2). \tag{13.25}$$

Here a is real analytic and $a \geq \frac{1}{2}$, so the surface $t = 0$ is non-characteristic. Analytic Cauchy data give a local analytic solution by Cauchy–Kowalevski; Holmgren uniqueness shows that a second weak solution with the same Cauchy data must agree with it locally.

The argument in the notes is a duality argument. Suppose u_1 and u_2 are solutions of the same Cauchy problem, with u_1 analytic and u_2 only in a weak class such as H^1_{loc}. Their difference

$$u = u_1 - u_2$$

satisfies $Lu = 0$ and has zero Cauchy data on B.

Let Q be a thin region between B and a nearby top hypersurface T. Green's formula has the form

$$\int_Q (Lu)\phi \, \mathrm{d}x \, \mathrm{d}t = \int_Q uL^*\phi \, \mathrm{d}x \, \mathrm{d}t + \int_B \mathcal{B}_B(u,\phi) + \int_T \mathcal{B}_T(u,\phi), \tag{13.26}$$

after localization removes lateral boundary terms. The zero Cauchy data for u annihilate the term on B.

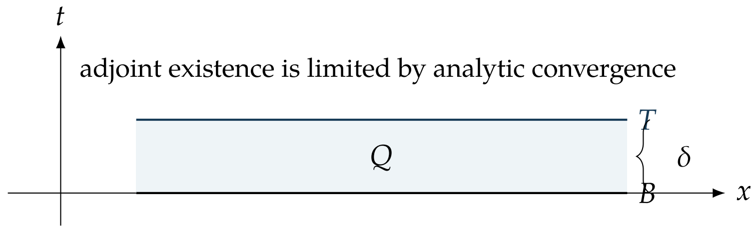
Choose ϕ by solving the backward adjoint Cauchy problem

$$L^*\phi = \psi, \qquad \partial_\nu^j \phi\big|_T = 0, \quad j = 0, \dots, m-1, \tag{13.27}$$

where ψ is analytic. Since T is non-characteristic, the Cauchy–Kowalevski theorem solves this problem in a sufficiently thin region; the thickness is limited by the radius of convergence of the analytic solution. The terminal data in (13.27) annihilate the term on T. Hence

$$0 = \int_Q (Lu)\phi = \int_Q u\psi.$$

The localized analytic duality argument then gives $u = 0$ near B. This is the mechanism behind Holmgren uniqueness for non-analytic solutions.



13.6 Symmetric hyperbolic systems

Many higher-order hyperbolic equations can be rewritten as first-order systems. After normalizing the positive time matrix, a symmetric hyperbolic system has the form

$$\partial_t u - \sum_{j=1}^{d} A^j(x,t)\partial_j u = 0, \qquad (A^j)^{\mathsf{T}} = A^j. \tag{13.28}$$

Assume that u is compactly supported in space, or decays sufficiently fast that no boundary term remains. Symmetry gives

$$\begin{aligned}\int_{\mathbb{R}^d} \left\langle A^j\partial_j u, u\right\rangle \,\mathrm{d}x &= -\int_{\mathbb{R}^d} \left\langle u, A^j\partial_j u\right\rangle \,\mathrm{d}x - \int_{\mathbb{R}^d} \left\langle (\partial_j A^j)u, u\right\rangle \,\mathrm{d}x \\ &= -\frac{1}{2}\int_{\mathbb{R}^d} \left\langle (\partial_j A^j)u, u\right\rangle \,\mathrm{d}x. \end{aligned} \tag{13.29}$$

Taking the L^2 inner product of (13.28) with u therefore yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\int_{\mathbb{R}^d} |u|^2 \,\mathrm{d}x\right) \le C\int_{\mathbb{R}^d} |u|^2 \,\mathrm{d}x. \tag{13.30}$$

Gronwall's inequality gives the energy estimate. Differentiating the system and repeating the same calculation gives higher-order estimates.

13.6.1 Characteristic hypersurfaces

Let $\phi(x,t) = c$ define a hypersurface, with conormal $(\partial_t\phi, \nabla_x\phi)$. For (13.28), it is characteristic when

$$\det\left((\partial_t\phi)I - \sum_{j=1}^{d} A^j\partial_j\phi\right) = 0. \tag{13.31}$$

Equivalently, with dual variables τ and ξ,

$$\det\left(\tau I - \sum_{j=1}^{d} A^j\xi_j\right) = 0. \tag{13.32}$$

For each (x,t,ξ), the matrix $\sum_j A^j\xi_j$ is symmetric. Hence its eigenvalues, the characteristic roots $\tau = \sigma_k(x,t,\xi)$, are real. Because the symbol is linear in ξ, these roots are positively homogeneous of degree one in ξ. This real characteristic geometry is the algebraic signature of a symmetric hyperbolic system.

13.7 Problems

Problem 13.1 (A quasilinear wave equation)**.** Let $a \in C^\infty(\mathbb{R})$ satisfy

$$a'(s) \geq 1 \qquad (s \in \mathbb{R}),$$

and consider

$$\partial_{tt} u - \partial_x \big(a(\partial_x u)\big) = 0 \qquad \text{on } [0,T] \times \mathbb{R}.$$

(a) Determine the type of the equation.
(b) For a smooth solution with sufficient decay at spatial infinity, derive the spatially differentiated energy estimate suggested by the equation.
(c) Consider the iteration

$$\partial_{tt} u^{n+1} - a'(\partial_x u^n)\partial_{xx} u^{n+1} = 0,$$

with common initial data. Under uniform bounds on the iterates sufficient for the energy calculation, show that the iteration is a contraction for sufficiently small T.

Solution. Since

$$\partial_x \big(a(\partial_x u)\big) = a'(\partial_x u)\partial_{xx} u,$$

the equation is

$$\partial_{tt} u - a'(\partial_x u)\partial_{xx} u = 0.$$

(a) The characteristic equation is

$$(\partial_t \phi)^2 - a'(\partial_x u)(\partial_x \phi)^2 = 0,$$

or

$$\partial_t \phi = \pm\sqrt{a'(\partial_x u)}\,\partial_x \phi.$$

Because $a'(\partial_x u) \geq 1$, the two characteristic roots are real and distinct. Thus the equation is strictly hyperbolic.

(b) Put

$$b = a'(\partial_x u), \qquad f_k = \partial_x^k \partial_t u, \qquad g_k = \partial_x^{k+1} u.$$

Differentiating the equation k times in x gives

$$\partial_t f_k - b\partial_x g_k = [\partial_x^k, b]\partial_{xx} u.$$

Multiplication by f_k yields the pointwise identity

$$\partial_t \left(\frac{1}{2} f_k^2 + \frac{1}{2} b g_k^2\right) - \partial_x (b f_k g_k) = \frac{1}{2}(\partial_t b) g_k^2 - (\partial_x b) f_k g_k + f_k [\partial_x^k, b]\partial_{xx} u.$$

Integrating over $\mathbb{R}$ makes the flux term vanish. Therefore, with

$$E_k(t) = \frac{1}{2}\int_{\mathbb{R}} \left(\left|\partial_x^k \partial_t u\right|^2 + a'(\partial_x u)\left|\partial_x^{k+1} u\right|^2\right) \mathrm{d}x,$$

one obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}E_k(t) = \int_{\mathbb{R}} \Big\{ \frac{1}{2} a''(\partial_x u)\partial_{xt} u \left|\partial_x^{k+1} u\right|^2 - a''(\partial_x u)\partial_{xx} u(\partial_x^k \partial_t u)(\partial_x^{k+1} u) + (\partial_x^k \partial_t u)[\partial_x^k, a'(\partial_x u)]\partial_{xx} u \Big\}\, \mathrm{d}x.$$

The mean-value theorem, the Sobolev embedding $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$, and the usual product estimates then give the schematic bound written in the calculation,

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}E_k(t)\right| \lesssim C_{\|\partial_x u\|_{L^\infty}} P(\|\partial_x u\|_{H^k}, \|\partial_t u\|_{H^k}),$$

for a polynomial P determined by k and by derivatives of a on the range of $\partial_x u$. Since $a' \geq 1$, E_k controls $\left\|\partial_x^k \partial_t u\right\|_{L^2}^2 + \left\|\partial_x^{k+1} u\right\|_{L^2}^2$.

(c) Define

$$w^n = u^{n+1} - u^n, \qquad b_n = a'(\partial_x u^n).$$

Subtracting two consecutive iteration equations gives

$$\partial_{tt} w^n - b_n \partial_{xx} w^n = (b_n - b_{n-1})\partial_{xx} u^n.$$

With

$$\mathcal{E}_n(t) = \frac{1}{2}\int_{\mathbb{R}} \left(|\partial_t w^n|^2 + b_n\, |\partial_x w^n|^2\right) \mathrm{d}x,$$

the same calculation as above gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_n(t) = \int_{\mathbb{R}} \Big\{ \frac{1}{2}(\partial_t b_n)\, |\partial_x w^n|^2 - (\partial_x b_n)(\partial_t w^n)(\partial_x w^n) + (b_n - b_{n-1})\partial_{xx} u^n \partial_t w^n \Big\}\, \mathrm{d}x.$$

If the iterates obey the uniform bounds used in the handwritten estimate, the mean-value theorem gives

$$|b_n - b_{n-1}| \leq C\left|\partial_x w^{n-1}\right|,$$

and hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_n(t) \leq C\mathcal{E}_n(t) + C\left\|\partial_x w^{n-1}(t)\right\|_{L^2}^2.$$

The differences have zero initial data. Since $b_n \geq 1$, integration in time and Gronwall's inequality imply

$$\sup_{0\leq t\leq T}\left(\|\partial_t w^n(t)\|_{L^2}^2 + \|\partial_x w^n(t)\|_{L^2}^2\right) \leq CTe^{CT} \sup_{0\leq t\leq T}\left\|\partial_x w^{n-1}(t)\right\|_{L^2}^2.$$

Choosing T so that $CTe^{CT} < 1$ makes the iteration a contraction. Thus $(\partial_t u^n, \partial_x u^n)$ converges in $C([0,T]; L^2(\mathbb{R}) \times L^2(\mathbb{R}))$, and the limit solves the equation in the corresponding energy class.

The function itself can be recovered from the common initial value by

$$u(t) = u(0) + \int_0^t \partial_s u(s)\,\mathrm{d}s.$$

If one additionally knows that $\partial_x u(t) \in L^1(\mathbb{R})$ and that $u(t,x) \to 0$ as $x \to -\infty$, then one may instead write

$$u(t,x) = \int_{-\infty}^{x} \partial_x u(t,s)\,\mathrm{d}s.$$

□

Problem 13.2 (Characteristic cones for the wave equation)**.** Determine the characteristic equation for the wave operator on $\mathbb{R}_t \times \mathbb{R}^d_x$ and identify the characteristic hypersurfaces through $(0, x_0)$.

Solution. A hypersurface $\phi(t,x) = 0$ is characteristic when

$$(\partial_t \phi)^2 = \sum_{i=1}^{d} (\partial_{x_i} \phi)^2.$$

The characteristic hypersurfaces through $(0, x_0)$ are the two sheets of the light cone,

$$t = \pm \left| x - x_0 \right|.$$

□

Problem 13.3 (Characteristics for a Hamilton–Jacobi equation)**.** For

$$\partial_t u = H(x, \nabla u), \qquad x \in \mathbb{R}^d,$$

derive the characteristic system for x and $p = \nabla u$.

Solution. Set $\phi = \nabla u$. Differentiating the equation gives

$$\partial_t \phi = \nabla_x H(x,\phi) + D_x\phi\, \nabla_p H(x,\phi).$$

Along a characteristic $x = x(t)$, put $p(t) = \phi(t, x(t))$. Choosing the curve so that the terms containing $D_x\phi$ cancel gives

$$\boxed{\frac{\mathrm{d}x}{\mathrm{d}t} = -\nabla_p H(x,p), \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = \nabla_x H(x,p).}$$

□

References

The following books were used as references in preparing these notes. The entries record the editions or printings consulted.

[1] Lawrence C. Evans, *Partial Differential Equations*, 1st ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.

[2] Gerald B. Folland, *Introduction to Partial Differential Equations*, 2nd ed., Princeton University Press, Princeton, NJ, 1995.

[3] Jürgen Jost, *Partial Differential Equations*, 1st ed., Graduate Texts in Mathematics, vol. 214, Springer-Verlag, New York, 2002.

[4] Jeffrey Rauch, *Partial Differential Equations*, Graduate Texts in Mathematics, vol. 128, Springer-Verlag, New York, 1991; corrected 2nd printing, 1997.